

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: March 26, 2004, 17:14:33 ; Search time 4233 seconds
(without alignments)
2855.770 Million cell updates/sec

Title: US-10-014-927-19
Perfect score: 1451
Sequence: 1 MMSRNRRTTVGNLPGDIRK.....RSKSRSRNSPVPSPVLSG 279

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US10014927/runat 24032004 152925 9896/app_query.fasta_1.455
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10014927 @CNC 1 1 3731 @runat 24032004 152925 9896 -NCPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.cm.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.roi.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1377.5	94.9	838	8	AV150486	AV150486 Arabidops
2	1272	87.7	762	6	AX505504	AX505504 Sequence
3	1032.5	71.2	933	8	AY050912	AY050912 Arabidops
4	904	62.3	59261	8	TI2M4	AC003114 Arabidops
5	891	61.4	5164	8	ATH131214	AJ131214 Arabidops
6	875	60.3	4044	6	AX040661	AX040661 Sequence
7	865	59.6	1229	8	AK118379	AK118379 Arabidops
8	865	59.6	1232	8	AY085920	AY085920 Arabidops
9	854.5	58.9	1523	8	AY056185	AY056185 Arabidops
10	854.5	58.9	2059	8	AK118074	AK118074 Arabidops
11	847	58.4	858	8	BT006316	BT006316 Arabidops
12	847	58.4	1540	8	AY128338	AY128338 Arabidops
13	846	58.3	1293	8	AK071503	AK071503 Oryza sat
14	839	57.8	1160	8	ATHRIENUCP	M98340 Arabidops
15	832.5	57.4	1352	8	AK106176	AK106176 Oryza sat
16	827.5	57.0	1238	8	AK061903	AK061903 Oryza sat
17	827.5	57.0	3430	8	AK121333	AK121333 Oryza sat
18	816	56.2	1899	8	AK072839	AK072839 Oryza sat
19	808	55.7	1501	8	AK065547	AK065547 Oryza sat
20	757	52.2	2196	8	AK103534	AK103534 Oryza sat
21	586.5	40.4	1618	5	BC042354	BC042354 Xenopus l
22	585	40.3	2369	5	BC046679	BC046679 Xenopus l
23	584.5	40.3	2765	6	BD189938	BD189938 Tumor ant
24	582.5	40.1	1428	6	BD094070	BD094070 Shear str
25	582.5	40.1	1428	9	HUMSF2P33	M69040 Human SF2p3
26	582.5	40.1	2708	9	BC010264	BC010264 Homo sapi
27	582.5	40.1	2765	9	AB062124	AB062124 Homo sapi
28	582.5	40.1	2785	10	BC058627	BC058627 Mus muscu
29	582.5	40.1	2878	10	BC046773	BC046773 Mus muscu
30	566.5	39.0	809	3	AF242767	AF242767 Caenorhab
31	566	39.0	2002	5	BC055511	BC055511 Dario rer
32	551.5	38.0	94536	8	TI4P8	AF069298 Arabidops
33	551.5	38.0	194892	8	ATCHRIV6	AL161494 Arabidops
34	550	37.9	1099	3	AK115991	AK115991 Ciona int
35	550	37.9	1391	3	AF232773	AF232773 Drosophil
36	550	37.9	1395	3	AY051919	AY051919 Drosophil
37	546.5	37.7	160415	2	AC140961	AC140961 Papio anu
38	543	37.4	1584	6	AX305513	AX305513 Sequence
39	543	37.4	1584	10	MWASF	X66091 M.musculus
40	542.5	37.4	1717	9	HUMASF	M72709 Human alter
41	542.5	37.4	2926	8	AF173640	AF173640 Arabidops
42	541.5	37.3	5285	6	AR204400	AR204400 Sequence
43	539.5	37.2	1860	5	BC056752	BC056752 Dario rer
44	525	36.2	1193	10	BC012217	BC012217 Mus muscu
45	520	35.8	257703	3	CEY111B2A	AL132904 Caenorhab

ALIGNMENTS

RESULT 1

AY150486 838 bp mRNA linear PLN 23-SEP-2002
LOCUS
DEFINITION Arabidopsis thaliana putative SF2/ASF splicing modulator Srp30
ACCESSION AY1509140 mRNA, complete cds.
VERSION AY150486
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (Chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 838)
Yanada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J.,
Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T.,
Davis,R.W., Ecker,J.R. and Theologis,A.
Arabidopsis Open Reading Frame (ORF) Clones
Unpublished
2 (bases 1 to 838)
Yanada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
Wu,H.C., Yu,G., Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J.,
Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tripp,M.G., Wu,T.,
Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.
FEATURES
source
1..838
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="1"
/clone="C105028"
/ecotype="Columbia"
/note="This clone is in PUNI 51."
1..838
/gene="Atlg09140"
1..807
/gene="Atlg09140"
/codon_start=1
/evidence=experimental
/product="putative SF2/ASF splicing modulator Srp30"
/protein_id="AA113011.1"
/db_xref="GI:23297698"
/translation="MSSRWRTIYVGNLPGDIRKCEVEDLFYKYGPIVDLDLKIPIPP
PGYAFVEFEDPRDADAIYGRDYGDFDGLRLVETAGRRFPSPVDRYSSYSASRA
PSRRDYVLVTGLPPASWQDLKHKAGVCFSEVPDRKMGSGVYDSYDDMK
YAIRKLDATFENAFSSAYIRVREYRSRSRSPDDSKYSRSRSGPSCSYSSKSR
SVSPARSISPRSRPLSRSLYSVRSQSRKSRSRSPVSPVISG"
808..838
/gene="Atlg09140"
3'UTR
ORIGIN
Alignment Scores:
Pred. No.: 5,09e-66 Length: 838
Score: 1377.50 Matches: 268
Percent Similarity: 96.06% Conservative: 0
Best Local Similarity: 96.06% Mismatches: 0
Query Match: 94.93% Indels: 11
DB: 8 Gaps: 1
US-10-014-927-19 (1-279) x AY150486 (1-838)
QY 1 MetSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
DB 1 ATGAGTAGCCGATGGATCGTACGATCTACGTTGGGATTTCCCTGGAGATTCGCGAG 60
QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40

61 TGTGAGTTGAAGATCTCTTCTACAAGTATGACCAATTTGTGACATGATTGTAAGATT 120
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
121 CCACCGAGACCTCTCTGGTTATGCCCTTTGTGAGATTTGAGATCTCTCGTGTGACACGAT 180
QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
181 GCATTTTATGAGACGATGATGTTATGATTTGATGGGTGTCGACTTCGGGTGGAGATTGCA 240
QY 81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSer 100
241 CATGGTGGTCGTAGATTTTTCACCACTCAGTTGATAGTACAGCAGCAGCTACAGTCGAGC 300
QY 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProSerAla 120
301 CGTGACCTTTCAAGACGCTCTGACTACCGCGTCTTGACCGGATTTACCGCTTCTGCT 360
QY 121 SerTyrGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
361 TCGTGGCAGGACCTTAAGGATCACATGCGAAAGCTGGAGATGTTCTGCTTCTCTGAAGTT 420
QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
421 TTCCCTGACCGTAAAGGCATGCTGGGGTGTGGATTTATAGCAACTATGATGATGAAG 480
QY 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
481 TAGCGCAATTAAGAAACTTGATGCCACTGAATTCGAAATGCTTTCTTAGTCTTATATA 540
QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
541 CGGTGAGGGAATATGATCGAGGAGTGTGAGTCAGCCAGCCAGATGATCTTAAGACTAT 600
QY 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
601 AGAAGCAGGAGTCGAGGCGGTGTCCTCAAGCTGTAGCTATAGTAGCAAGCAGCAGGAGTGT 560
QY 221 SerProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeuTyr 240
661 TCACCTGTGTAGTCCATTTCCCGCGGTTCACGCGCCCTTAGTCTGCTCGCTCGCTATAC 720
QY 241 SerSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTyrIleSerGlnSerArg 260
721 AGCTCTGTCTCAAGG-----TCCCAATCAAGA 747
QY 261 SerLysSerArgSerArgSerArgSerArgSerProValSerProValIleSerGly 279
748 TCAAAATCAAGATCAAGATCAAGATCGAATTTCCAGTTTCCAGTTTGATGATCTGCT 804
RESULT 2
AX506504
LOCUS
DEFINITION Sequence 1199 from Patent WO0216655.
ACCESSION AX506504
VERSION AX506504.1 GI:23387741
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1
Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
Patent: WO 0216655-A 1199 28-FEB-2002;
The Scripps Research Institute (US) ; Syngenta Participations AG
(CH)
FEATURES
Location/Qualifiers
source
1..762
/organism="Arabidopsis thaliana"

ORIGIN		/mol_type="unassigned DNA" /db_xref="taxon:3702"	
Alignment Scores:			
Pred. No.:	2.17e-60	Length:	762
Score:	1272.00	Matches:	253
Percent Similarity:	91.67%	Conservative:	0
Best Local Similarity:	91.67%	Mismatches:	1
Query Match:	87.66%	Indels:	23
DB:	6	Gaps:	2
US-10-014-927-19 (1-279) x AX506504 (1-762)			
Qy	1	MetSerSerArgTrrPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys	20
Db	1	ATGAGTAGCCGATCGAATCGTACGATCTACGTTGGGAATTTGCTCGAGATATTCGCAAG	60
Qy	21	CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle	40
Db	61	TGTGAGGTTGAAGATCTCTTCTCAAGTATGGAACCAATTGTGGACATTGATTGAAGATT	120
Qy	41	ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspArg	60
Db	121	CCACCGAGACTCTCGTTATGCTTTGTGCGATTTGAAGATCCTCGTATGACGACGAT	180
Qy	61	AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla	80
Db	181	GCAATTTATGGACGTGATGTTATGATTTGTGCGGTGCGATTTGGTTGAGATTGCA	240
Qy	81	HisGlyGlyArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSer	100
Db	241	CATGGTGGTGGTATTTTCCACATCAGTTGATAGGTACAGCAGCAGTACAGTGGGAGC	300
Qy	101	ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla	120
Db	301	CGTGACCTTCAAGACGCTCTGACTACCGCGTCTGTGACCGGATTTACCGCTTCTGCT	360
Qy	121	SerTrrPglNAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal	140
Db	361	TCGTGGCAGACCTTAAGATCATCATCGCGCAAGCTGGAGATGCTGCTTCTGAGTT	420
Qy	141	PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys	160
Db	421	TTCCCTGACCGTAAAGGCATGCTCGGGTTGTGGATTATAGCAACTATGATGATGAAG	480
Qy	161	TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle	180
Db	481	TACGCAATAAGAAACTTGTATGCTGCACTGAAATTCGAAATGCTTCTCTAGTCTTATA	540
Qy	181	ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr	200
Db	541	CGGGTAGGGGAATATGATGCGAGGTGTGAGTCGAAGCCGACGATGATTTCTAAAGCTAT	600
Qy	201	ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal	220
Db	601	AGAAGCAGGAGTGG-----AGCCGTGT-GTG	626
Qy	221	SerProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeuTyr	240
Db	627	TCACCTGTAGATCCATTTCCCGCGTTCACGCCCTTAGTGTCTGCTCGCTCGCTATAC	686
Qy	241	SerSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrrPlesSerGlnSerArg	260
Db	687	AGCTCTGTCTCAAGG-----TCCCAATCAAGA	713
Qy	261	SerLysSerArgSerArgSerArgSerAsnSerProValSerProVal	276
Db	714	TCAAAATCAAGATCAAGATCAAGATCGAATTTCTCCAGTTTCCACCTGTA	761
RESULT 3			
AY050912			
LOCUS			
		933 bp	mrna linear PLN 18-SEP-2002

Arabidopsis thaliana putative SP2/ASF splicing modulator Srp30 protein (At1g09140) mRNA, partial cds.

AY050912.1 GI:15292956

FLI CDNA.

Arabidopsis thaliana (thale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

1 (bases 1 to 933)

Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kaniya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tracy,S.E., Shinozaki,K., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

Arabidopsis Full Length cDNA Clones

Unpublished

2 (bases 1 to 933)

Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Banno,F., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kaniya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.

Direct Submission

Submitted (24-JUL-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kaniya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada,K., Liu,S.X., Pham,P.K., Banh,J., Banno,F., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Yamamura,Y., Yu,G., Yu,S., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tracy,S.E., Davis,R.W., Ecker,J.R. and Theologis,A.

Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis Genome submitted to GenBank.

Location/Qualifiers

1..933

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="1"

/clone="RAPL08-12-105 (R11038)"

/ecotype="Columbia"

/note="This clone is in a modified pBluescript vector (FLC-1) as a BamHI/XhoI insert."

<1..933

/gene="At1g09140"

<1..626

/gene="At1g09140"

/codon_start=3

/evidence=experimental

/product="putative SP2/ASF splicing modulator Srp30 protein"


/protein_id="AAK93589.2"

```
/db_xref="GI:21954085"
/translation="IYGRDGYDFDGCRLRVIAHGRRFSPVDYRYSYSSASRAPSR
RSDYLVLTGLPSSAQWQDKHMRKAGDYCFSEVFPDRKMSGVDPYSNDYDMKYAI
RKDLATERNAFSAIYIRVEYRSVSRSPDDSKFSRSGSRGSCSYSSKRSRVS
PARSISPRSLRSLRSLYSSVSRSQSRSKRSTRSGNSPVSPVISG"
misc_difference 1
/gene="Atlg09140"
/note="compared to genomic sequence"
/replaces="c"
misc_difference 148
/gene="Atlg09140"
/note="compared to genomic sequence"
/replaces="g"
misc_difference 579
/gene="Atlg09140"
/note="compared to genomic sequence resulting in an amino
acid sequence difference"
/replaces="c"
627. 933
/gene="Atlg09140"
3'UTR
ORIGIN
Alignment Scores:
Pred. No.: 2,09e-47 Length: 933
Score: 1032.50 Matches: 205
Percent Similarity: 94.50% Conservative: 1
Best Local Similarity: 94.04% Mismatches: 1
Query Match: 71.16% Indels: 11
DB: 8 Gaps: 1
US-10-014-927-19 (1-279) x AY050912 (1-933)
QY 62 IletyRGlyArgAspGlyTyrAspPheAspGlyCysArgLeuA:GValGluileAlaHis 81
DB 3 ATTTATGACGTGATGGTTATGATTTGATGGTGTGCACCTTCGGTTGAGATTCACAT 62
QY 82 GlyGlyArgGlySerProSerValAspArgTyrSerSerTyrSerAlaSerArg 101
DB 63 GSTGTGCTAGATTTCACATCAGTTGATAGTACAGCAGCAGTACAGTGCAGCGT 122
QY 102 AlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProSerAlaSer 121
DB 123 GCACCTTCAAGACGCTCTGACTACACCGTGTGTGACCGGATTACCGCTTCTGCTTCG 182
QY 122 TrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPhe 141
DB 183 TGGCAGGACCTTAAGGATACATGCGCAAGCTGGAGATGCTGCTTCTTGAAGTTTC 242
QY 142 ProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyr 161
DB 243 CTGTACCGTAAGGACATGCTGGGGTTGTGGATTATAGCACTATGATGATGATGATAC 302
QY 162 AlalieArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrileArg 181
DB 303 GCAATAAGGAACTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 362
QY 182 ValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArg 201
DB 363 GTGAGGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
QY 202 SerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSer 221
DB 423 ACAGAGATCGGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 482
QY 222 ProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeuTyrSer 241
DB 483 CTGTAGATCATTTCCCGGTTTCAGGCGGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCT 542
QY 242 SerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSer 261
DB 543 TCTGTCTCAAGG-----TCCCAATCAAGTCA 569
QY 262 LysSerArgSerArgSerArgSerAsnSerProValSerProValIleSerGly 279
```

```
Db 570 AATCAAGAACAGATCAAGATCCAGTTCCTCCAGTTTACCTGTGATATCGT 623
RESULT 4
T12M4
LOCUS
DEFINITION
T12M4 59261 bp DNA linear PLN 23-JUN-1998
Arabidopsis thaliana chromosome 1 BAC T12M4 sequence, complete
sequence.
ACCESSION
AC003114
VERSION
AC003114.1 GI:3249094
KEYWORDS
HTG.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 59261)
Vysotskaia V.S., Schwartz J.R., Toriumi M., Yu G., Oji, O.,
Kwan, A., Liu, S., Li, J., Araujo, R., Au, M., Brendel, V., Buehler, E.,
Conway, A.B., Conway, A.R., Dewar, K., Feng, J., Kim, C., Kurtz, D.,
Li, Y., Palm, C.J., Shinn, P., Sun, H., Davis, R.W., Ecker, J.R.,
Federspiel, N.A. and Theologis, A.
Arabidopsis thaliana chromosome 1 BAC T12M4 sequence, complete
sequence
Unpublished (1998)
This sequence is of BAC T12M4 from Arabidopsis thaliana chromosome
1. The sequence does not represent the sequence of the entire
insert of this clone. It is shorter by 11040 bp because we submit
only the unique sequence of the clone. However, in order to
facilitate the joining of overlapping clones in the future for
creation of larger contigs, we provide a small overlap (200 bp)
between overlapping submitted clones. The 3' end of this sequence
overlaps by 200 bp the 5' end of the sequence of the BAC F7G19.
2 (bases 1 to 59261)
Theologis, A.
Direct Submission
Submitted (24-NOV-1997) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
3 (bases 1 to 59261)
Theologis, A.
Direct Submission
Submitted (23-JAN-1998) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
4 (bases 1 to 59261)
Theologis, A.
Direct Submission
Submitted (23-JUN-1998) Plant Gene Expression Center, 800 Buchanan
St., Albany, CA 94710, USA
On Jun 23, 1998 this sequence version replaced gi:2804593.
The sequence of BAC T12M4 from Arabidopsis thaliana chromosome 1.
FEATURES
source
1. .59261
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/clone="T12M4"
996. .4040
/gene="T12M4.1"
join(996. .1187,1532. .1693,2159. .2221,2271. .2494,
2581. .2662,2741. .2823,3006. .3045,3122. .3221,3334. .4040)
/gene="T12M4.1"
/note="contains similarity to dihydrofolate reductase
(dfr1) gb|L13703 from Schizosaccharomyces pombe. ESTs
gb|N37567 and gb|43002 come from this gene."
/codon_start=1
/evidence=not experimental
/protein_id="AAC24078.1"
/db_xref="GI:3249095"
/translation="MGTSSCGDHEKQRIEDSQVLLYYKTSVPDLDLVSVFESS
CNSLGLGRVLRSPKGVNTVGGKLTALFEHTAAKSNCLFEGTDFKLASCHPLNDK
VAEECGTSLSRVVEKFMQLGSIILQSFISRCRVPLITPPCNSAEHLDDGDKSENK
```


QY	21	CysGluValGluAspLeuPheTyrIys	29
Db	56491	TGTGAGGTGAAGATCTCTCTACAGGTTTGAAAATTCCTCTTTCTCTCGATAAAAT	56550
QY	29		29
Db	56551	TGAATTCATTATGACTTGTGTTGGTTCATAAATTTGCAATTCCTCTGCTGAGACAATT	56610
QY	30		39
Db	56611	TAAATCGACTCTTATGATATTTGTTTCAGTATGAGCAATTTGGACATTTGATTGAG	56670
QY	40	IleProProArgProGlyTyrAlaPheValGlu	51
Db	56671	ATTCCACCGAGACCTCTCTGTTATGCTTTGTCGA-GGTATATTGATCAAGTACAAATTT	56729
QY	51		51
Db	56730	GTTTTTTTCTCTCTTGTAAATAGTATAGGCTAATGACTAAGATGTTTGTATTGTGG	56789
QY	52	PheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAspPhe	70
Db	56790	CAGTTTGAAGATCTCTGATGTCAGACGATCAATTTATGACGTGATGTTATGATTTT	56849
QY	71	AspGlyCysArgLeuArg	76
Db	56850	GATGGTGTGACTTCGGTTAGTAAACGCATGATGAAGCTAGCTTAATTTCTGTAAT	56909
QY	77		84
Db	56910	TTCCTGTAAGGTGTTATCTTTTGTGTGATGTTTTAGGTTGAGATTGCACATGTTGTCG	56969
QY	84	GArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSerArgAlaProse	104
Db	56970	TAGATTTTCCATCAGTTGATAGGTACAGCAGCAGCTACAGTCGAGCCGTGCACCTTC	57029
QY	104	rArgArgSerAspTyrArg	110
Db	57030	AAGACGCTCTGACTACCGGGTTGTAGAGCTTCTCGATTGTTGTTATTTGGTGTGTGT	57089
QY	111		116
Db	57090	AAAATTTTATTTTGAAAACCTCAATTTTACTACCTAAACATGATGCTTCTGACCGGAT	57149
QY	116	euProProSerAlaSerTrpGlnAspLeuIys	126
Db	57150	TACCGCCCTCTGCTTCGTGGCAGGACCTTAA-GGTAAAGGACACTATATAGTCTTTTCT	57208
QY	126		126
Db	57209	CTGAATGTTGGTCTCTATATCANGTTTGGATTATCTCTTTTCTGAAATGATGTTATT	57268
QY	127		140
Db	57269	TCCTATTTACGGGTGATTAGGATCACATGCGCAAGCTGGAGATGCTGCTTCTCTGAG	57328
QY	140	alPheProAspArgIys	145
Db	57329	TTTTCCCGTGAAGGTGAGTGTGACATTCGATAGTTTGGATAAGCTTTTGTATGATG	57388
QY	145		145
Db	57389	TGTTAGTAAATPAGTCTTTGTGAAGGAGAATAGGTGTTAAGCATCTGAACCTGCTAAACTC	57448
QY	146		159
Db	57449	ACATTTCAGTATTTCTTTGTAGGCATGCTGGGGTTGTGGATTATAGCAACTATGATGATA	57508
QY	159	etLysTyrAla	162
Db	57509	TGAAGTACGC-AGTAAGTTTATATCTTTTGCAACCAAAATGTTCTCTGACTTATGCCTTA	57567

QY	162		162
Db	57568	GACTGCTTTTGTTCATAGTATACCGAGCTGAATTTATCTTCTCGAGCCAGTGTGGA	57627
QY	163		171
Db	57628	TCTTTGATGTTCCCTTAAATTTTTCATGTCGACAGATAGGAACTTGTGCACTGAT	57687
QY	171	heArgAsnAlaPheSerSerAlaTyrIleArg	181
Db	57688	TTTGGAAATGCTTCTCTAGTCTTATATACGGGTATGTTGTTATGCTTCTTTGATTTTG	57747
QY	182		186
Db	57748	TTAAGCATAAAGTGGATATGGAGTCATCTCTGAATTTACTGTTTCAGGTGAGGAATATGAG	57807
QY	187	SerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSer	206
Db	57808	TCGAGAGTGTGAGTCGAAGCCAGATGATTTCTAAAGCTATAGAACAGAGTGGAGC	57867
QY	207	ArgGlyProSerCysSerTyrSerLysSerArg	218
Db	57868	CGTGGTCCAGCTGATGCTATAGTAGCAAGACGAGGCTCTTTTTTTTTTTTTTTTTT	57927
QY	218		218
Db	57928	CATAAACCTTAGACATATAAGGATTTTATTGTAATTTATTGAAATTAACCTGACTT	57987
QY	219		234
Db	57988	CYAAATGCAATCAGGTGTCCACCTGCTAGTAGTCCATTTTCCCGCGTTCACGCCCTTAG	58047
QY	234	rArgSerArgSerLeuTyrSerSerValSerArg 245	
Db	58048	TGTTTCTGCTCGCTATACAGCTCTGTCTCAAGG 58081	



RESULT 5	
LOCUS	ATH131214 5164 bp DNA linear PLN 07-MAY-1999
DEFINITION	Arabidopsis thaliana srp30 gene, exons 1-12.
ACCESSION	AF131214
VERSION	AF131214.1 GI:4775269
KEYWORDS	SF2/ASF-like splicing modulator; srp30 gene.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE	1
AUTHORS	Lopato,S., Kalyna,M., Dörner,S., Kobayashi,R., Krainer,A.R. and Barta,A.
TITLE	atSrp30, one of two SF2/ASF-like proteins from Arabidopsis thaliana, regulates splicing of specific plant genes
JOURNAL	Genes Dev. 13 (8), 987-1001 (1999)
MEDLINE	99234087
PUBMED	10215626
REFERENCE	2 (bases 1 to 5164)
AUTHORS	Barta,A.
TITLE	Direct Submission
JOURNAL	Submitted (02-DEC-1998) Barta A., Institute for Biochemistry, University of Vienna, Dr. Bohrgasse 9/3,, A-1030 Vienna, AUSTRIA
FEATURES	Location/Qualifiers
source	1..5164
	/organism="Arabidopsis thaliana"
	/mol_type="genomic DNA"
	/db_xref="taxon:3702"
mRNA	join(1806..2014,2138..2203,2290..2364,2444..2546,
	2632..2678,2786..2843,2968..3017,3161..3217,3291..3403,
exon	3501..3578,4515..4572,4844..5125)
	1806..2014
	/number=1
gene	1928..4858
	/gene="srp30"



join(1928.. .2014,.2138.. .2203,.2290.. .2364,.2444.. .2546,
 2632.. .2678,.2786.. .2843,.2968.. .3017,.3161.. .3217,.3291.. .3403),
 3501.. .3578,.4515.. .4572,.4844.. .4858)
 /gene="srp30"
 /codon_start=1
 /product="SF2/ASF-like splicing modulator Srp30"
 /protein_id="CAB42557.1"
 /db_xref="GI:4775270"
 /db_xref="GOA:Q9XFR5"
 /db_xref="SPTREMBL:O9XFR5"
 /translation="MSSRNRTIYVGNLPGRIRKEVEDLFYKYGPIVDILKIPPRP
 PGVAFEFEDRDADDAIYGRDGYDFDCGLRLVEIAHGRRFPSPDRYSSSYASRA
 PSRSDDVLVTGCLPPSAWQLDKHMRKAGDVCFSEVPDRKGMSGVWDYSNYDDMK
 YATKLDATEFNFAFSAAIVRVEYESVSRSRDDSKSVESRSRGSGCSYSSKS
 SVSPARSISRPSPLRSLSRSLYSSVRSQRSRKSRSRNSPSPVPSIG"
 join(1928.. .2014,.2138.. .2203,.2290.. .2364,.2444.. .2546,
 2632.. .2678,.2786.. .2843,.2968.. .3017,.3161.. .3217,.3291.. .3403),
 3501.. .3578,.4178.. .4214)
 /gene="srp30"
 /note="alternative exon 10a"
 /codon_start=1
 /product="SF2/ASF-like splicing modulator Srp30, variant
 1"
 /protein_id="CAB42558.1"
 /db_xref="GI:4775271"
 /db_xref="GOA:Q9XFR6"
 /db_xref="SPTREMBL:O9XFR6"
 /translation="MSSRNRTIYVGNLPGRIRKEVEDLFYKYGPIVDILKIPPRP
 PGVAFEFEDRDADDAIYGRDGYDFDCGLRLVEIAHGRRFPSPDRYSSSYASRA
 PSRSDDVLVTGCLPPSAWQLDKHMRKAGDVCFSEVPDRKGMSGVWDYSNYDDMK
 YATKLDATEFNFAFSAAIVRVEYESVSRSRDDSKSVESRSRGSGCSYSSKS
 SVSPARISRPSPLRSLSRSLYSSVRSQRSRKSRSGSLLRAGDWI"
 2015.. .2137
 /gene="srp30"
 /number=1
 2138.. .2203
 /gene="srp30"
 /number=2
 2204.. .2289
 /gene="srp30"
 /number=2
 2290.. .2364
 /gene="srp30"
 /number=3
 2365.. .2443
 /gene="srp30"
 /number=3
 2444.. .2546
 /gene="srp30"
 /number=4
 2547.. .2631
 /gene="srp30"
 /number=4
 2632.. .2678
 /gene="srp30"
 /number=5
 2679.. .2785
 /gene="srp30"
 /number=5
 2786.. .2843
 /gene="srp30"
 /number=6
 2844.. .2967
 /gene="srp30"
 /number=6
 2968.. .3017
 /gene="srp30"
 /number=7
 3018.. .3160
 /gene="srp30"
 /number=7
 3161.. .3217
 /gene="srp30"

Db 2347 GATGGTGTGCGACTTCGGGTAGTAAGCATGATGAAGCTAGCTTAATTTCTGTAAT 2406
Qy 77 -----ValGluLeuAlaHisGlyGly 84
Db 2407 TTCCTGTAAGGTGTATCTTTGTGTGATGTTTTAGGTTGAGATTGCACATGCTGGTCG 2466
Qy 84 GArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSerArgAlaProSe 104
Db 2467 TAGATTTTACCATCAGTTGATAGGTACAGCAGCAGCTACAGTCCGACCGCTGCACCTTC 2526
Qy 104 rArgArgSerAspTyrArg ----- 110
Db 2527 AAGACGCTCTGACTACCGCGGTTGTAGAGTCCTTCGATGTGTATTTGGTGTGTGT 2586
Qy 111 -----ValLeuValThrGlyL 116
Db 2587 AAAATTTTATATTTGAAACCTCATTTTACTACCTAAACATGATGCTTGTGACCGGAT 2646
Qy 116 euProProSerAlaSerTyrGlnAspLeuLys ----- 126
Db 2647 TACCGCCTTCTGCTTCGTGGCAGGACCTTAA-GGTAAGGGGACACTATATAGTCTTTTCT 2705
Qy 126 ----- 126
Db 2706 CTGAATGTGTCTCTATATCATGTTTGGATTATCTCTTTCTGAAATGATGTTAT 2765
Qy 127 -----AspHisMetArgLysAlaGlyAspValCysPheSerGluV 140
Db 2766 TGCTATTTACGGGTGATAGGATCACATGCGCAAGCTGGAGATGCTGCTTCTCTGAAG 2825
Qy 140 alPheProAspArg-Lys ----- 145
Db 2826 TTTTCCCTGACCGTAAGGTGAGTGACATTCGATAGTTTGGATAAGCTTTTGTGATTGAT 2885
Qy 145 ----- 145
Db 2886 GTGTTAGTAATTAAGTCTTTCTGAAGGAGATAGGTGTTAAGCATCTGAATCTGCTAACT 2945
Qy 146 -----GlyMetSerGlyValValAspTyrSerAsnTyrAspAsp 158
Db 2946 CACATTCAGTATTTCTTTGTAGGATGTCGGGTGTGGGATATAGCAACTATGATGAT 3005
Qy 159 MetLysTyrAla ----- 162
Db 3006 ATGAGTACGC-AGTAAGTTTATATCTTTGCAACGCAATGTTCTCGGACTTATGCTT 3064
Qy 162 ----- 162
Db 3065 AGACTGCTTTGTTTCATAGTATACCGAGCTGAATTTATCTTCCTCGAGGCCAGTGTGG 3124
Qy 163 -----lleArgLysLeuAspAlaThrGlu 170
Db 3125 ATCTTTGATGTTCCCTTAAATTTTGTATGTGACAGATAAGGAAACTTGATGCCACTGAA 3184
Qy 171 PheArgAsnAlaPheSerSerAlaTyrIleArg ----- 181
Db 3185 TTTTCAAATGCTTTCTCTAGTCTTATATACGGTATGTTGTATGCTTTCTTTGATTTT 3244
Qy 182 -----ValArgGluTyrG1 186
Db 3245 GTTAAGCATAGTGATGATGATCATCTCTGAATTTACTGTTTCAGGTGAGGGAATATGA 3304
Qy 186 uSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSe 206
Db 3305 GTCGAGGAGTGTGAGTCCGAAGCCAGATGATTTCTAAAAGCTATAGAAGCAGGAGTCGGAG 3364
Qy 206 rArgGlyProSerCysSerTyrSerSerLysSerArg ----- 218
Db 3365 CCGTGGTCCAAGCTGTAGCTATAGTAGCAAGCAGCAGGAGGCTTTTCTTTTCTTTTCTTTT 3424
Qy 218 ----- 218
Db 3425 TCATAAACCTAAGACATATAAGGGATTTTATTTGTAACCTATTATGAAATTAACCTGACT 3484

Qy 219 -----SerValSerProAlaArgSerIleSerProArgSerArgProLeuS 234
Db 3485 TCTAATGCAATGACAGTGTGTACCTGTAGATCAATTTCCCGCGTTTCACGGCCCTTA 3544
Qy 234 exArgSerArgSerLeuTyrSerSerValSerArg 245
Db 3545 GTGCGTTCGCTCGCTATACAGCTCTGTCTCAAG 3579
RESULT 6
AX040661 4044 bp DNA linear PAT 23-NOV-2000
LOCUS Sequence 18 from Patent WO0065059.
DEFINITION AX040661
ACCESSION AX040661
VERSION AX040661.1 GI:11340389
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1
AUTHORS Barta,A., Lopato,S., Kalyna,M. and Dörner,S.
TITLE Splicing factor
JOURNAL Patent: WO 0065059-A 18 02-NOV-2000;
(sterreichisches Forschungszentrum Seibersdorf Ges.m.b.H.; (AT)
FEATURES
Location/Qualifiers
1..4044
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="Beschreibung des unbekannten ORGANISM:Genom
atSRp30"
ORIGIN
Alignment Scores:
Pred. No.: 3,268-38 Length: 4044
Score: 875.00 Matches: 242
Percent Similarity: 43.53% Conservative: 0
Best Local Similarity: 43.53% Mismatches: 3
Query Match: 60.30% Indels: 314
DB: Gaps: 9
US-10-014-927-19 (1-279) x AX040661 (1-4044)
Qy 1 MetSerSerArgTyrAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
Db 808 ATGAGTAGCCGATGGATCGATCGATACGATACGTTGGGAAATTTGCCCTGGAGATATTCGCAAG 867
Qy 21 CysGluValGluAspLeuPheTyrLys ----- 29
Db 868 TGTGAGGTTGAAGATCTCTCTTACAAAGGTTTGAAAATTTCTCTTTTCTCTCGATAAAA 927
Qy 29 ----- 29
Db 928 TTGAATTCATATAGACTAGTTGGGTTCAATAATTTGCAATTCCTCTTCTGCTGAGACAAAT 987
Qy 30 -----TyrGlyProIleValAspIleAspLeuLys 39
Db 988 TTAATTCGACTCTTATGATATATTTGTTTCAGTATGACCAATTTGTCGACATTGATTGAA 1047
Qy 39 sIleProProArgProProGlyTyrAlaPheValGlu ----- 51
Db 1048 GATTCACCGAGACCTCTCTGTTATGCTTTCGA-GGTATATTCATCAAGTACAAAT 1106
Qy 51 ----- 51
Db 1107 TGTTTTTTCTTCTTTGTAATAGTAGGCTAATAGACTAGATAGTATTGTTTATTGGTG 1166
Qy 52 -----PheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyrAspPh 70
Db 1167 GCAGTTTGAAGATCTCTCGTGATGCAGACGATCAATTTTATGACGCTGATGTTTATGATT 1226
Qy 70 exaspGlyCysArgLeuArg ----- 76

Db 1227 TGATGGGTGCGACTTCGGGTAGTAGTAACCGCATGATGAAGAGCTAGCTTAATTTCTGTAA 1286
Qy 77 -----ValGluIleAlaHisGlyGlyA 84
Db 1287 TTTCTTGTAAAGGTGTTATCTTTCTGTGATGTTTTAGGTTGAGATTGCACATGGTGGTC 1346
Qy 84 IGAArgPheSerProSerValAspArgTyrSerSerTySerAlaSerArgAlaProS 104
Db 1347 GTAGATTTCACCATCAGTTGATAGGTACAGCAGCAGCTACAGTGGGAGCGGTGCACCTT 1406
Qy 104 eArgArgSerAspTyrArg----- 110
Db 1407 CAAGACGCTCTGACTACCGGGTTGTAGAGTCTTCTCGATTGTTGTTATTTGGTGTGTG 1466
Qy 111 -----ValLeuValThrGly 115
Db 1467 TAAATTTTATATTGAAACTCATTTTACTACCTAAACATGATGCTGTGACCGGA 1526
Qy 116 LeuProSerAlaSerTyrGlnAspLeuLys----- 126
Db 1527 TTACCGCTTCTGCTTCTGCGCAGACCTTAA-GGTAAGGACACTATATAGTCTTTTTC 1585
Qy 126 ----- 126
Db 1586 TCTGAATGTGGTCTCTATATCATGATGTTTGGATTATCTCTTTCTGAAATGATCTAT 1645
Qy 127 -----AspHisMetArgIleAlaGlyAspValCysPheSerGlu 139
Db 1646 TTGCTATTTAGGGGTGATTAGGATCATCGGCAAGCTGGAGATGCTGCTTCTCTGAA 1705
Qy 140 ValPheProAspArg-Lys----- 145
Db 1706 GTTTTCCTCGACCGTAAGGTGAGTTCGACATTCGATGTTGGTAAGCTTTTGTATTGA 1765
Qy 145 ----- 145
Db 1766 TGTGTAGTAAATPAGTCTTTGTGAAGGAGATAGGTGTTAAGCATCTGAACCTGCTAAAC 1825
Qy 146 -----GlyMetSerGlyValValAspTyrSerAsnTyrAspAs 158
Db 1826 TCACATTTCAGTATTTCTTTAGGCATGCTCTGGGTTGTGGATTATAGCAACTATGATGA 1885
Qy 158 pMetLysTyrAla----- 162
Db 1886 TATGAATGAGC-AGTAAGTTTATATCTTTGCAACGCAATGTTCTGCGATTATGCTT 1944
Qy 162 ----- 162
Db 1945 TAGACTGCTTTGTTTCATAGTATACCGAGCTGAATTTATCTTCTGGAGCGCCAGGTGG 2004
Qy 163 -----IleArgLysLeuAspAlaThrGly 170
Db 2005 GATCTTTGATGTTCCCTTAAATTTTGTATGTGACAGATAAGGAACTTGATGCCACTGA 2064
Qy 170 uPheArgAsnAlaPheSerSerAlaTyrIleArg----- 181
Db 2065 ATTTTCGAATGCTTTCTTAGTGCTTATACGGGTATGTTGTTATGTCCTTTCTTTGATT 2124
Qy 182 -----ValArgGluTyrG 186
Db 2125 TGTTAAGCATAGTGATATGGAGTCATCTCTGAAATTTACTGTTTCAGGTGAGGGAATAG 2184
Qy 186 luSerArgSerValSerArgSerProAspAspSerTySerTyArgSerArgSerArgS 206
Db 2185 AGTCGAGAGTGTGAGTCGAAGCCAGACATGATTTCAAAGCTATAGAAGCAGAGTCGGA 2244
Qy 206 eArgGlyProSerCysSerTySerSerLysSerArg----- 218
Db 2245 GCCGTGTCCTCAAGCTGTAGCTATAGTAGCAAGAGCAGAGGCTCTTTTTTTTTTTTT 2304
Qy 218 ----- 218
Db 2305 TTTCAAAACCTAAGACATATAAGGGATTTTATTTGTAACCTATTATGAATAAATGAC 2364

Qy 219 -----SerValSerProAlaArgSerIleSerProArgSerArgProLeu 233
Db 2365 TTTCAATATGCAATGCAATGCTGCTACCTGCTAGATCCATTTCCCGCGTTCACGCCCTT 2424
Qy 234 SerArgSerArgSerLeuTyrSerSerValSerArg 245
Db 2425 AGTCGTTCTGCTGCTATACAGTCTGCTCTCAAGG 2460
RESULT 7
AKI18379 1229 bp mRNA linear PLN 06-DEC-2002
LOCUS Arabidopsis thaliana Atlg02840 mRNA for putative ribonucleoprotein
DEFINITION SF-2, complete cds, clone: RAFL19-64-H16.
VERSION AKI18379
KEYWORDS FLI CDNA; CAP trapper.
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
Kawai, J., Hayashizaki, Y. and Shinozaki, K.
TITLE Arabidopsis thaliana full-length cDNA
JOURNAL Published Only in Database (2002)
REFERENCE 2 (bases 1 to 1229)
AUTHORS Seki, M., Iida, K., Satou, M., Sakurai, T., Akiyama, K., Ishida, J.,
Nakajima, M., Enju, A., Kamiya, A., Narusaka, M., Carninci, P.,
Kawai, J., Hayashizaki, Y. and Shinozaki, K.
TITLE Direct Submission
JOURNAL Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences
Center; 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa
230-0045, Japan (E-mail: mseki@gsc.riken.go.jp,
URL: http://pfweb.gsc.riken.go.jp. Tel: 81-45-503-9625,
Fax: 81-45-503-9586)
COMMENT An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al. (1998) Plant J. 15:707-720;
Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FLC-1-E vector (Carninci et
al. (2001) Genomics 77:79-90) digested with BamHI and SalI.
This clone is in a modified pBluescript vector.
Please visit our web site (http://pfweb.gsc.riken.go.jp/) for
further details.
FEATURES
Location/Qualifiers
1..1229
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="1"
/clone="RAFL19-64-H16"
/note="Common name: thale cress"
1..1229
gene /gene="Atlg02840/F22D16_30"
93..1004
CDS /gene="Atlg02840/F22D16_30"
/codon_start=1
/product="putative ribonucleoprotein SF-2"
/protein_id="BAC42991.1"
/db_xref="GI:26451794"
ORIGIN
Alignment Scores: 2.93e-38 Length: 1229
Pred. No.:

US-10-014-927-19 (1-279) x AY056185 (1-1523)	
QY	1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
Db	49 ATGACACCGGTCGAGTAGAACGATTTACGTCGGAACCTTCCTCCGCGATATCCGTGAA 108
QY	21 CysGluValGluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuLys 40
Db	109 AGAGAAGTTGAAGACTTCTTCAGTAAGTATGAGCTGTGTTCATTCGATTTGAAGATT 168
QY	41 PropCoArgProProGlyTyrAlaPheValGluPheGluAspProCoArgAspAlaAsp 60
Db	169 CCGCCGAGGCTCCAGGCTATGCATTCGTGAGTTTGAGGATGCTGTAAGCTGATGAT 228
QY	61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuA-gValGluIleAla 80
Db	229 GCAATTTATGCGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 288
QY	81 HisGlyGlyArgAspPheSerProSerValAspArgTyrSerSerSerTyrSerAla --- 99
Db	289 CATGCTGGAGCGGTTTCATCATGATGCA-----CGCGGTAGTTATAGTGGTTCGT 339
QY	100 -----SerArgAlaProSerArgArg 106
Db	340 GGTCTGTGGCGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 399
QY	107 SerAspTyrArgValLeuThrGlyLeuProProSerAlaSerTrpGlnAspLeuLys 126
Db	400 TCAGAGTACCGGTTGTAGTGCAGTTTGTCTCACTGCTGCTGCTGCTGCTGCTGCTGCT 459
QY	127 AspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGly 146
Db	460 GATCATCATCGTAAAGGAGGAGAGGTTGTTTCTCAAGTGTTCGTGATGTTAGAGGT 519
QY	147 MetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIle-ArgLysLe 166
Db	520 ACACTGGAATTTGTAGATTATACCACTACGAGGACATGAAATATGCGATAAAAAAGCT 579
QY	166 uaspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGl 186
Db	580 CGATGACACAGAGTTTCGAATGCGTTTCTCATGATATGTTTCGGTTTAGAATATGA 639
QY	186 uSerArgSerValSerArgSerProAspAspSerLysSerTyr---ArgSerArgSerAr 205
Db	640 TTCAGAGGAGGATTCGAGGAGCCCGCGCGTGAAGATCTTCTTAAGACCCGACCGC 699
QY	205 gSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSe 225
Db	700 TGGACGGAGTCCAGC-----CGTAGTCGTAGCCGACAGCAGAGCAGGAGCAAG 750
QY	225 rIleSerProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerAr 245
Db	751 CAGAAAGTCAAGGCTAAATCTTTCGTAGATCGCTGCA-----AAATCTACATCGAG 804
QY	245 gSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSerLysSerArgSe 265
Db	805 AICT-----CCTGCTCTCTGCTCCGCTC 828
QY	265 rArgSerArgSerAsnSerPro 272
Db	829 TAAAGTCGAGGTCACTGCTCTCCA 850
RESULT 10	
AK118074	
LOCUS	
DEFINITION	
Arabidopsis thaliana At4g02430 mRNA for unknown protein, complete cds, clone: RAFL19-32-J05.	
ACCESSION	
AK118074.1	
VERSION	
AK118074.1 GI:26451202	
KEYWORDS	
FLI CDNA; CAP trapper.	
SOURCE	
Arabidopsis thaliana (thale cress)	
ORGANISM	
Arabidopsis thaliana	
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.	
1 Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J., Nakajima,M., Enju,A., Kamiya,A., Narusaka,M., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.	
Arabidopsis thaliana full-length cDNA	
Published Only in Database (2002)	
2 (bases 1 to 2059)	
Seki,M., Iida,K., Satou,M., Sakurai,T., Akiyama,K., Ishida,J., Nakajima,M., Enju,A., Kamiya,A., Narusaka,M., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.	
Direct Submission	
Submitted (25-NOV-2002) Motoaki Seki, RIKEN Genomic Sciences Center; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:mseki@gs.riken.go.jp, URL:http://pfweb.gsc.riken.go.jp, Tel:81-45-503-9625, Fax:81-45-503-9586)	
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al. (1998) Plant J. 15:707-720; Seki et al. (2002) Science 296:141-145). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda PLC-1-E vector (Carninci et al. (2001) Genomics 77:79-90) digested with BamHI and SalI. This clone is in a modified pBluescript vector.	
Please visit our web site (http://pfweb.gsc.riken.go.jp/) for further details.	
FEATURES	
source	
1..2059	
/organism="Arabidopsis thaliana"	
/mol_type="mRNA"	
/cultivar="Columbia"	
/db_xref="taxon:3702"	
/chromosome="4"	
/clone="RAFL19-32-J05"	
/note="common name: thale cress"	
gene	
1..2059	
/gene="At4g02430/Tl4P8_21"	
153..689	
/gene="At4g02430/Tl4P8_21"	
/codon_start=1	
/product="unknown protein"	
/protein_id="BAC42705.1"	
/db_xref="GI:26451203"	
/translation="MSSRSSTIYVGNLPGDIRREVEDLFSKYGVQIDLUKIPRP PGVAFEPEDADDADAYRGDYPDGHHLRLAHLGRRSHDARSYSGRGGGR GGCGGGRGSRSEYRVVVSGLPSSASWDLKDHMRKGSEVCFQVFRDGRGTTG IVDYTSYEDMKYALKAR"	
ORIGIN	
Alignment Scores:	
Pred. No.: 1-93e-37 Length: 2059	
Score: 854.50 Matches: 181	
Percent Similarity: 72.92% Conservative: 29	
Best Local Similarity: 62.85% Mismatches: 42	
Query Match: 58.89% Indels: 36	
DB: 6 Gaps: 6	
US-10-014-927-19 (1-279) x AK118074 (1-2059)	
QY	1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
Db	153 ATGACACCGGTCGAGTAGAACGATTTACGTCGGAACCTTCCTCCGCGATATCCGTGAA 212
QY	21 CysGluValGluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuLys 40
Db	213 AGAGAAGTTGAAGACTTGTTCAGTAAGTATGAGCTGTGTTCATTCGATTTGAAGATT 272
QY	41 PropCoArgProProGlyTyrAlaPheValGluPheGluAspProCoArgAspAlaAsp 60
Db	273 CCGCCGAGGCTCCAGGCTATGCATTCGTGAGTTTGAGGATGCTGTAAGCTGATGAT 332
QY	61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuA-gValGluIleAla 80

Qy 61 AlalleTyArgAspGlyTyArgPheAspGlyCysArgLeuArgValGluIleAla 80
 Db 181 GCATTTCATGGTCTGTGATGCTATGACTTGTGATGGCATGTTTGGAGGTGGAAATGGCG 240
 Qy 81 HisGlyGlyArgArgPheSerProSerValAspArgTyArgSerSerTyArgSer 98
 Db 241 CATGCTGGAGGCGT-----TCATCAGATGATACCTCGGGGTAGTTTCAATGCTGGT 291
 Qy 99 -----AlaSerArgAlaProSerArgArgSer 107
 Db 292 GGCCGTGGTGGTGGTGGCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 351
 Qy 108 AspTyArgValLeuValThrGlyLeuProProSerAlaSerTyrGlnAspLeuLysAsp 127
 Db 352 GAGTTTGGTGTCTAGTCACAGGCTTGCCTTCATCTCTCTTGGCAGATCTCAAGGAT 411
 Qy 128 HisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMet 147
 Db 412 CACATCGGTAAGGAGGCGATGCTGTTCTCGCAAGTGTACCGTGATGCTAGAGGACA 471
 Qy 148 SerGlyValValAspTyArgSerAsnTyArgAspMetLysTyrAlaIleArgLysLeuAsp 167
 Db 472 ACTGGAGTTGTTGATTACCTGCTATGAGGACATGAAGTATCGCTGAAAAGCTCGAC 531
 Qy 168 AlaThrGluPheArgAsnAlaPheSerSerAlaTyIleArgValArgGluTyArgLys 187
 Db 532 GACACAGAGTTTCGAATGCGTTTTCGAATGGATATGTCGGGTAGAGAAATATGATCA 591
 Qy 188 ArgSerValSerArgSerProAspAspSerLysSerTyr---ArgSerArgSerArgSer 206
 Db 592 AGGAAGGATTTAGAGTCTAGCGGGGAGATCCTATTCTAAGACCGCAGCCGACG 651
 Qy 207 ArgGlyProSerCysSerTyArgSerSerLysSerArgSerValSerProAlaArgSerIle 226
 Db 652 CGTGACGACGAGCGTG-----AGCGAAGCAGGAGCAGCAGGAGCAGGAGCAGA 702
 Qy 227 SerProArgSerArgProLeuSerArgSerArgSerLeuTyArgSerSerValSerArgSer 246
 Db 703 AGTCCCAAGGCAAG-----TCTTACGTAGGTCCCTCGAAATCTACATCAAGATCT 756
 Qy 247 GlySerLeuLeuArgAlaGlyAspTyrIleSerGlnSerArgSerLysSerArgSerArg 266
 Db 757 CTGGC-----CCCCGCTCGAG 774
 Qy 267 SerArgSerAsnSerPro 272
 Db 775 TCAGGTACCGTCTCCA 792
 RESULT 12
 AY128338
 LOCUS
 DEFINITION
 Arabidopsis thaliana SF2/ASF-like splicing modulator Srp30,
 putative (Atg02840) mRNA, complete cds.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (bases 1 to 1540)
 Tripp, M., Southwick, A., Karlin-Neumann, G., Nguyen, M., Miranda, M.,
 Palin, C.J., Bowser, L., Jones, T., Banh, J., Carninci, P., Chen, H.,
 Cheuk, R., Chung, M.K., Hayashizaki, Y., Ishida, J., Kamiya, A.,
 Kawai, J., Kim, C., Lin, J., Liu, S.X., Narusaka, M., Pham, P.K.,
 Sakano, H., Sakurai, T., Satou, M., Seki, M., Shinn, P., Yamada, K.,
 Shinozaki, K., Ecker, J., Theologis, A. and Davis, R.W.
 Direct Submission
 Submitted (01-JUL-2002) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA

COMMENT

e-mail for correspondence: arab@sequence.stanford.edu

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y., and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Tripp, M., Nguyen, M., Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palin, C.J., Bowser, L., Jones, T., Banh, J., Chen, H., Cheuk, R., Chung, M.K., Kim, C., Lin, J., Liu, S.X., Pham, P.K., Sakano, H., Shinn, P., Yamada, K., Ecker, J., Theologis, A. and Davis, R.W.

Tripp, M. (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Davis, R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES

Location/Qualifiers
 source

1..1540
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clones="RAFL04-19-E10 (R15578)"
 /ecotype="Columbia"
 /note="This clone is in pBluescript"

gene

1..1540
 /gene="Atlg02840"
 /note="synonym: F22D16.16"

CDS

49..906
 /gene="Atlg02840"
 /codon_start=1
 /product="SF2/ASF-like splicing modulator Srp30, putative"
 /note="synonym: F22D16.16"
 /protein_id="AA091541.1"
 /db_xref="GI:22135918"
 /translation="MSSRSRTVYVGNLPGDIRREVEDLFSKGVVQIDLVPPRP
 PCYAFVDDADAEADITHGRDGYDFGHRRLVLAHGRSSDDTRGSFNGGSGG
 RGRGSGRSPSRSEPRVLTGLPSSAWODLKHMRKGDVCFQVYRDARGTTGV
 VDTCYEMKALKLDDTEFRNAFSNGYVAVREYDSRSDSRGSRGSRSRSR
 GRSVRSRSRSRSRSRSPAKSRSRSPAKSRSRSPAKSRSRSPAKSRSRSPAKSRSR
 DHNTIISGFL"

ORIGIN

Alignment Scores:
 Pred. No.: 3,528-37 Length: 1540
 Score: 847.00 Matches: 177
 Percent Similarity: 73.08% Conservative: 32
 Best Local Similarity: 61.89% Mismatches: 41
 Query Match: 58.37% Indels: 36
 DB: Gaps: 6

US-10-014-927-19 (1-279) x AY128338 (1-1540)

Qy 1 MetSerArgTyrAsnArgThrIleTyValGlyValGluLeuProGlyValAspIleArgLys 20
 Db 49 ATGAGCGTCGTCGAGTAGAACCCTGTCGGAACCTTCCTGGGATATCGTGAG 108
 Qy 21 CysGluValGluAspLeuPheTyArgTyArgTyArgTyArgTyArgTyArgTyArgTyArg 40
 Db 109 AGAGAGTCAAGATTGTTGTTAGTAAAGTATGACCTGTGTTCAAAATTCATTGAAGTT 168
 Qy 41 ProProArgProProGlyTyArgAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
 Db 169 CCTCCAAGCCCTCCTGGTTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTAT 228
 Qy 61 AlaIleTyGlyArgAspGlyTyArgPheAspGlyCysArgLeuArgValGluIleAla 80
 Db 229 GCTATTCAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 288
 Qy 81 HisGlyGlyArgArgPheSerProSerValAspArgTyArgSerSerSerTyArgSer 98
 Db 289 CATGGTGGAGGCGT-----ICATCAGATGATCTCGGGGTAGTTTCAATGCTGGT 339

JOURNAL	JOURNAL	TITLE	COMMENT
Science 301 (5631), 376-379 (2003)	2752273	Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of	
2 (bases 1 to 1293)	12869764	Agrobiological Sciences, Department of Molecular Genetics, Head of	
REFERENCE		Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki	
AUTHORS		305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp,	
		Tel:81-29-838-7007, Fax:81-29-838-7007]	
		This clone is one of the 28K full-length cDNA clones from japonica	
		rice.	
		URL : http://cdna01.dna.affrc.go.jp/cDNA/	
		NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,	
		Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,	
		Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,	
		Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and	
		Yamamoto, M.	
		FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,	
		Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,	
		Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,	
		Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,	
		Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,	
		Yoshimura, A., Matsubara, K. and Murakami, K.	
		Genome Exploration Research Group in Riken Genomic Sciences Center	
		and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,	
		Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,	
		Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,	
		Hirooka, T., Hori, F., Iida, J., Imamura, K., Inotani, K., Ishii, Y.,	
		Itoh, M., Kagawa, I., Kanagawa, S., Kondo, H., Kawai, J.,	
		Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,	
		Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,	
		Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,	
		Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,	
		Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,	
		Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,	
		Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,	
		Yasunishi, A. and Hayashizaki, Y.	
FEATURES		Location/Qualifiers	
source		1. .1293	
		/organism="Oryza sativa (japonica cultivar-group)"	
		/mol_type="mRNA"	
		/cultivar="Nipponbare"	
		/db_xref="taxon:39947"	
		/clone="J02309IN08"	
ORIGIN			
Alignment Scores:			
Pred. No.:	3 276-37	Length:	1293
Score:	846.00	Matches:	186
Percent Similarity:	75.27%	Conservative:	24

Db 502 ACTGGAGTGTGTTGATTACACCTGCTATGAGGACATGAAGTATGCGGTGAAAAAGCTCGAC 561
Qy 168 AlAThrGluPheArgAanAlaPheSerSerAlaTyIleArgValArgGluTyrGluSer 187
Db 562 GACACAGATTTCGAATGCGTTTCGAATGGATATGTCGCGGTAGAGAAATGATGATCA 621
Qy 188 ArgSerValSerArgSerProAspAspSerLysSerTyr---ArgSerArgSerArgSer 206
Db 522 AGGAGGATTTCTAGGAGTCTAGCCGGGGAAGATCTCTATCTAAGAGCGCGACGCCACG 681
Qy 207 ArgGlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSerIle 226
Db 682 CGTGGACGAAGCGTG-----AGCGAAGCAGGAGCAGAGAGGAGGAGCAGACAGA 732
Qy 227 SerProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerArgSer 246
Db 733 AGTCCCAAGCAAG-----TCTTACGAGTAGTCCCTCCCAAAATCTCATCATCAAGATCT 786
Qy 247 GlySerLeuLeuAlaGlyAspTyrIleSerGlnSerArgSerLysSer-----Arg 264
Db 787 CTGGGCCCCCGC-----TCGAGTCAAGGTCACCGTCTCCAGAGA 828
Qy 265 SerArgSerArgSerAanSerProValSerProVal 276
Db 829 TCGGTTCAAGATCAAGATCTCTCTACCTTCTGTT 864
AK106176 1352 bp mRNA linear PLN 24-JUL-2003
Oryza sativa (japonica cultivar-group) cDNA clone:001-208-C08, full
insert sequence.
ACCESSION AK106176
VERSION AK106176.1 GI:32991385
KEYWORDS F11 cDNA; oligo capping.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Superfamily: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1
REFERENCE
AUTHORS
The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team,
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group: Otono, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Miura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
JOURNAL Science 301 (5631), 376-379 (2003)
MEDLINE 22752273
PUBMED 12869764
REFERENCE
AUTHORS
2 (bases 1 to 1352)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, J., Iida, Y., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kawagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,

Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Otono, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Tagiyama, A., Takahashi, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.
Direct Submission
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
Tel:81-29-838-7007, Fax:81-29-838-7007)
This clone is one of the 28K full-length cDNA clones from japonica
rice.
URL : http://cdna01.dna.affrc.go.jp/cDNA/
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and
Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, R., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.
FEATURES
Location/Qualifiers
1..1352
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="001-208-C08"
ORIGIN
Alignment Scores:
Pred. No.: 1.83e-36 Length: 1352
Score: 832.50 Matches: 184
Percent Similarity: 72.11% Conservative: 28
Best Local Similarity: 62.53% Mismatches: 57
Query Match: 57.37% Indels: 25
DB: 8 Gaps: 7
US-10-014-927-19 (1-279) x AK106176 (1-1352)
Qy 1 MetSerArgTyrAsnArgThrIleTyrValGlyAanLeuProGlyAspIleArgLys 20
Db 145 ATTCAGTAGCGCTGGAGCAGGACAAATTTATGTGCGGAACCTCCCGGCGACATCGAGG 204
Qy 21 CysGluValGluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuLysIle 40
Db 205 AGGAGGAGTAGAGATCTGTCTTACAGGATGACCCATTTGTTGATTTGATCTTGAAGATC 264
Qy 41 ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60

```

Db      265  CCCCCAAGCGCGCTGGTTTGGCTTTGGTTGAGTTCGAGATCCCGTGATGCTGAAGAT 324
Qy      61  AlalIeTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuValGluIleAla 80
Db      325  GCAATCCGTGACGGGATGGCTATAATTTTATGGGAACCGTTTGAGGGTGGAACTTCCA 384
Qy      81  HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSer 100
Db      385  CATGGTGGGAGAGTAAGTCTCTCTCC-----TTTAACTATCTGGTGGAGGGGGA 435
Qy      101  Arg-----AlaProSerArgSerAspTyrArgValLeuValThrGlyLeuProPro 118
Db      436  CGTCGTGGGGGTGTATCTAGGCATACAGAGTATCGCGTCTGTGTACTGGACTACCTTCT 495
Qy      119  SerIaSerTrpGlnAspLeuLeuAspHisMetArgLysAlaGlyAspValCysPheSer 138
Db      496  TCAGCATCATGGCAAGACCTAAAGACCATATGGAAATGCTGGTGATGTTTGTACTCT 555
Qy      139  GluValPheProAspArgLysGlyMetSerClyValValAspTyrSerAsnTyrAspAsp 158
Db      556  GAAGTGTACCGTGGGGGTGGTGGTACTATAGGAATTGTAGACTACACAACTATGATGAC 615
Qy      159  MetLysTyrAlaIleAtcLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAla 178
Db      616  ATGAAGTATGCTATCAGGAACACTTGTGATTTCTGATTTAAGATGCTTCTCAAGGCC 675
Qy      179  TyrIleArgValArgGluTyrGlu-----SerArgSerValSerArgSerProAsp 195
Db      676  TATATAGGGTGAAGGATGATGATGGAAGCGCAGCGTTCATACTCAAGAGCGGAAGC 735
Qy      196  AspSer-----LysSerTyr---ArgSerArgSerArgSerArgGlyProSerCysSer 212
Db      736  CGAAGTCGAGGTCGCAGCTACAGCAGAAGCAGAGAGTCCAAAGCAAGTCTCCCAAGGAAAG 795
Qy      213  TyrSerSerLysSerArgSerValSerProIaIaArgSerIleSerProArgSerArgPro 232
Db      796  TCATCACCGCCTTCAGCATCTAGATCTGATCGAGATCAGCTTCTCCCGTTCAGATCA 855
Qy      233  LeuSerArgSerArgSerLeuTyrSerSerValSerArgSerGlySerLeuLeuArgAla 252
Db      856  GAGTCAAAGGGGGCGCTCTCCATCAAGATCACAGCAAGATCTCAGTCTCCAAATACTTCT 915
Qy      253  -----GlyAspTrpIleSer-----Gln 258
Db      916  CCTGCCAATGGTGTGATGCAGCAAGCCCCCAAGAAACGACGCCCAAGCAGGAGCCCCCAAG 975
Qy      259  SerArgSerLysSerArgSerArgSerArgSerArgSerArgSerPro 272
Db      976  AAACGTAGCCCAAGCAGGAGGCCCATCAGTTCCCGATCTCTCT 1017

```

Search completed: March 26, 2004, 19:37:55
Job time : 4277 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 26, 2004, 17:13:02 ; Search time 491 Seconds
(without alignments)
2413.946 Million cell updates/sec

Title: US-10-014-927-19
Perfect score: 1451
Sequence: 1 MSSRWNTIYGNLPGDIRK.....RSKRSRSRNSPVPVWSG 279

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+ p2n.model -DEV=xlh
-Q/can2_1/USPTO spool/US10014927/runat 24032004_152925_9886/app_query.fasta_1.455
-DB=N_Geneseq 25Jan04 -Qfmt=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10014927 @CN_1_1_470 @runat 24032004_152925_9886 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-FEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOF=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq 29Jan04: *
1: Geneseqn1980s: *
2: Geneseqn1990s: *
3: Geneseqn2000s: *
4: Geneseqn2001as: *
5: Geneseqn2001bs: *
6: Geneseqn2002as: *
7: Geneseqn2003as: *
8: Geneseqn2003bs: *
9: Geneseqn2003cs: *
10: Geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1359.5	93.7	925	3 AAC49248	AAC49248 Arabidops
2	1359.5	93.7	1213	3 AAC40180	AAC40180 Arabidops
3	1343	92.6	771	3 AAC49242	AAC49242 Arabidops
4	1272	87.7	762	6 AB213394	AB213394 Arabidops
5	1159	79.9	714	3 AAC42865	AAC42865 Arabidops
6	875	60.3	4043	3 AAC81899	AAC81899 A. thalia
7	865	59.6	1232	3 AAC34814	AAC34814 Arabidops
8	863	59.5	1190	3 AAC38528	AAC38528 Arabidops

9	838.5	57.8	885	3 AAC42654	AAC42654 Arabidops
10	805	55.5	1241	3 AAC37026	AAC37026 Arabidops
11	584.5	40.3	2765	9 ADD22452	ADD22452 HLA-B*46 T
12	582.5	40.1	1428	4 AAH02904	AAH02904 Human she
13	582.5	40.1	1428	7 AAL60055	AAL60055 Human PC0
14	550	37.9	1478	4 ABL29431	ABL29431 Drosophil
15	543	37.4	1584	6 ABI99368	ABI99368 Mouse isc
16	541.5	37.3	5285	2 AAV69915	AAV69915 Arabidops
17	537.5	37.0	498	3 AAC44527	AAC44527 Zea may
18	517	35.6	1069	6 ABN96845	ABN96845 Gene #334
19	517	35.6	1162	3 AAF16226	AAF16226 Human pro
20	517	35.6	1218	5 ABV30279	ABV30279 Human pro
21	517	35.6	1475	6 ABQ54393	ABQ54393 Human ova
22	472	32.5	479	3 AAC35783	AAC35783 Zea may
23	471	32.5	495	6 ABI99907	ABI99907 Mouse isc
24	457.5	31.5	568	3 AAC34694	AAC34694 Arabidops
25	456	31.4	3948	4 ABL29430	ABL29430 Drosophil
26	440.5	30.4	3697	7 ABX62942	ABX62942 Human act
27	436	30.0	289	3 AAC49252	AAC49252 Arabidops
28	432.5	29.8	545	7 ABX57250	ABX57250 Arabidops
29	431	29.7	1644	4 ABL30487	ABL30487 Drosophil
30	431	29.7	2076	6 ABK83839	ABK83839 Human cdn
31	418.5	28.8	1454	6 ABK84288	ABK84288 Human cdn
32	418.5	28.8	1534	7 ABX63268	ABX63268 Human pan
33	416.5	28.7	2212	3 AAC98914	AAC98914 Human pan
34	412.5	28.4	759	6 ABX77425	ABX77425 Frog embr
35	407.5	28.1	446	6 ABV94679	ABV94679 Human pan
36	402.5	27.7	727	5 ABV16009	ABV16009 Human pro
37	391	26.9	492	3 AAC47771	AAC47771 Zea may
38	388	26.7	345	6 ABQ85913	ABQ85913 Arabidops
39	376	25.9	3872	5 AAS64831	AAS64831 DNA encod
40	374.5	25.8	471	3 AAC41641	AAC41641 Arabidops
41	374	25.8	444	5 ABV45807	ABV45807 Human pro
42	344	23.7	746	6 ABX77297	ABX77297 Frog embr
43	325	22.4	489	8 ACH35932	ACH35932 Human end
44	298	20.5	1110	9 ADD01227	ADD01227 Human nuc
45	297	20.5	2367	2 AAZ41956	AAZ41956 Human myo

ALIGNMENTS

RESULT 1
AAC49248
ID AAC49248 standard; DNA; 925 BP.
XX AAC49248;
AC AAC49248;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 60471.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
FN EF1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.

19-JUL-1999,	PR	99US-0144335P,	PR
20-JUL-1999,	PR	99US-01444352P,	PR
20-JUL-1999,	PR	99US-01446322P,	PR
21-JUL-1999,	PR	99US-0144884P,	PR
21-JUL-1999,	PR	99US-0144814P,	PR
21-JUL-1999,	PR	99US-0145086P,	PR
21-JUL-1999,	PR	99US-0145088P,	PR
22-JUL-1999,	PR	99US-0145085P,	PR
22-JUL-1999,	PR	99US-0145087P,	PR
22-JUL-1999,	PR	99US-0145089P,	PR
22-JUL-1999,	PR	99US-0145132P,	PR
23-JUL-1999,	PR	99US-0145114P,	PR
23-JUL-1999,	PR	99US-0145218P,	PR
26-JUL-1999,	PR	99US-0145224P,	PR
26-JUL-1999,	PR	99US-0145276P,	PR
27-JUL-1999,	PR	99US-0145913P,	PR
27-JUL-1999,	PR	99US-0145918P,	PR
27-JUL-1999,	PR	99US-0145919P,	PR
28-JUL-1999,	PR	99US-0145951P,	PR
02-AUG-1999,	PR	99US-0146386P,	PR
02-AUG-1999,	PR	99US-0146388P,	PR
02-AUG-1999,	PR	99US-0146389P,	PR
03-AUG-1999,	PR	99US-0147038P,	PR
04-AUG-1999,	PR	99US-0147204P,	PR
04-AUG-1999,	PR	99US-0147302P,	PR
05-AUG-1999,	PR	99US-0147192P,	PR
05-AUG-1999,	PR	99US-0147260P,	PR
06-AUG-1999,	PR	99US-0147303P,	PR
06-AUG-1999,	PR	99US-0147416P,	PR
09-AUG-1999,	PR	99US-0147493P,	PR
09-AUG-1999,	PR	99US-01479935P,	PR
10-AUG-1999,	PR	99US-0148171P,	PR
11-AUG-1999,	PR	99US-0148331P,	PR
12-AUG-1999,	PR	99US-01483341P,	PR
13-AUG-1999,	PR	99US-0148565P,	PR
13-AUG-1999,	PR	99US-0148684P,	PR
16-AUG-1999,	PR	99US-0149368P,	PR
17-AUG-1999,	PR	99US-0149175P,	PR
17-AUG-1999,	PR	99US-0149428P,	PR
20-AUG-1999,	PR	99US-0149723P,	PR
20-AUG-1999,	PR	99US-0149929P,	PR
20-AUG-1999,	PR	99US-0149902P,	PR
23-AUG-1999,	PR	99US-0149930P,	PR
25-AUG-1999,	PR	99US-0150566P,	PR
25-AUG-1999,	PR	99US-0150684P,	PR
27-AUG-1999,	PR	99US-0151065P,	PR
27-AUG-1999,	PR	99US-0151066P,	PR
27-AUG-1999,	PR	99US-0151080P,	PR
30-AUG-1999,	PR	99US-0151303P,	PR
31-AUG-1999,	PR	99US-0151438P,	PR
01-SEP-1999,	PR	99US-0151930P,	PR
01-SEP-1999,	PR	99US-0152363P,	PR
10-SEP-1999,	PR	99US-0153070P,	PR
13-SEP-1999,	PR	99US-0153758P,	PR
13-SEP-1999,	PR	99US-0154018P,	PR
16-SEP-1999,	PR	99US-0154039P,	PR
16-SEP-1999,	PR	99US-0154779P,	PR
22-SEP-1999,	PR	99US-0155139P,	PR
23-SEP-1999,	PR	99US-0155486P,	PR
23-SEP-1999,	PR	99US-0156598P,	PR
24-SEP-1999,	PR	99US-0156598P,	PR
28-SEP-1999,	PR	99US-0156596P,	PR
04-OCT-1999,	PR	99US-0157117P,	PR
05-OCT-1999,	PR	99US-0157753P,	PR
06-OCT-1999,	PR	99US-0157865P,	PR
07-OCT-1999,	PR	99US-0158028P,	PR
08-OCT-1999,	PR	99US-0158232P,	PR
12-OCT-1999,	PR	99US-0158369P,	PR
13-OCT-1999,	PR	99US-0159293P,	PR
13-OCT-1999,	PR	99US-0159294P,	PR
13-OCT-1999,	PR	99US-0159295P,	PR
14-OCT-1999,	PR	99US-0159329P,	PR

PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 14-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 17-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 24-JUN-1999; 99US-0140696P.
PR 28-JUN-1999; 99US-0142823P.
PR 29-JUN-1999; 99US-0142830P.
PR 30-JUN-1999; 99US-0142831P.
PR 01-JUL-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142058P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144003P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144633P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146387P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157733P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159844P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.

```
PR 29-OCT-1999; 99US-0162142P.
Alignment Scores:
Pred. No.: 3,88e-94 Length: 1213
Score: 1359.50 Matches: 267
Percent Similarity: 95.02% Conservative: 0
Best Local Similarity: 95.02% Mismatches: 1
Query Match: 93.69% Indels: 13
DB: 3 Gaps: 2
US-10-014-927-19 (1-279) x AAC40180 (1-1213)
QY 1 MetSerSerArgTyrAsnArgThrIleTyrValGlyValAsnLeuProGlyAspIleArgLys 20
Db 134 ATGAGTAGCGGATGGAATCGTACGATCTACGTTGGGAATTTGCTGGAGATATTCGCAAG 193
QY 21 CysGluValGluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuLysIle 40
Db 194 TGTGAGTTGAAGATCTCTTACAAATGATGACCAATTTGTGACATTTGAATTTGAAGATT 253
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
Db 254 CCACCGAGACCTCTCGTTTATCGCTTTGTCGAGTTTGAAGATCTCTCGTATGCGAGACAT 313
QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db 314 GCAATTTATGGACGATGATGATTTTGAAGGTGTCGACTTCGGGTTGAGATTGCA 373
QY 81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSer 100
Db 374 CATGGTGGTCTAGATTTTACCACATCAGTTGATAGTACAGCAGCAGCTACAGTCCGAGC 433
QY 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
Db 434 CGTGACCTTCAAGACGCTCTGACTATCGCTGCTGTGACCGGATTACCGCTTCTGCT 493
QY 121 SerTyrGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
Db 494 TCGTGCAGGACCTTAAGGATCACATGCGCAAGCTGGAGATGTCTGCTTCTCGAAGTT 553
QY 141 PheProAspArgGlyGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
Db 554 TTCCTTGACCGTAAAGCAATGCTGGGGTTGGATTATAGCACTATGATGATATGAG 613
QY 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
Db 614 TAGCATAAAGGAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
QY 181 ArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
Db 674 CGGGTCAGGAGATATGAGTCGAGGAGTGTGAGTTCGAAGCCAGATGATTTCTAAAGCTAT 733
QY 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
Db 734 AGAAGCAGGAGTCGAGCGCGTGTCCAAAGCTGAGCTATAGTAGCAAGCAGGAGTGTG 793
QY 221 SerProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeuTyr 240
Db 794 TCACCTGCTAGATCCCATTTCCCGCGCTTACGGCCCTTAGTCGCTTCGCTCGCCATAC 853
QY 241 SerSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArg 260
Db 854 AGCTCTGTCTCAAGG-----TCCCAATCAAGA 880
QY 261 SerLysSerArgSer-----ArgSerArgSerAsnSerProValSerProValIleSer 278
Db 881 TCAAAATCAAGATCAACAYCAGCATCAAGATCGAATTTCTCCAGTTTCCAGCTGTGATATCT 940
QY 279 Gly 279
Db 941 GGA 943
RESULT 3
```

```
AAC49242
ID AAC49242 standard; DNA; 771 BP.
AC AAC49242;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 60449.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 16-APR-1999; 99US-0128714P.
PR 19-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
```

PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-01399750P.	PR	27-AUG-1999;	99US-0151085P.
PR	18-JUN-1999;	99US-01399763P.	PR	27-AUG-1999;	99US-0151086P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	22-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	24-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	28-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	29-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	30-JUN-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142330P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142603P.	PR	24-SEP-1999;	99US-0156458P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156596P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0158293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144332P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145152P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145152P.	PR	21-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	22-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	25-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	28-JUL-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	02-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149369P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			

Alignment Scores:
Pred. No.: 4,06e-93
Score: 1343.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 92.56%
DB: 3

Length:
Matches: 771
Conservative: 256
Mismatch: 0
Indels: 0
Gaps: 0

US-10-014-927-19 (1-279) x AAC49242 (1-771)

Qy	1	MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys	20
Db	1	ATGAGTAGCCGATGGAAATCGTACGATCTACGTTGGAAATTCCTGGAGATATTCGAAG	60
Qy	21	CysGluValGluAspLeuPheTyrIleValGlyProIleValAspIleAspLeuLysIle	40
Db	61	TGTGAGGTTGAGATCTCTTACAAAGTATGGACCAATTTGTGACATTCATTGAAGATT	120

QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
DB 121 CCACGAGACCTCTCGTTATGCGCTTTGTCAGATTGAAGATCTCTGATGACAGCAT 180
QY 61 AlaileTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
DB 181 GCATTTATGACGTGATGTTATGATTTGATGGGTGTCACCTTCGGTTGAGATTGCA 240
QY 81 HisGlyGlyArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSer 100
DB 241 CATGTGCTCGTAGATTTTACCATCATGTTAGTACAGTACAGCAGCAGTACAGTGCAGC 300
QY 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
DB 301 CGTGACCTTCAAGACGCTCTGACTACCGCTGCTGTGTGACCGGATTCACGCTTCGCT 360
QY 121 SerTyrGlnAspLeuLysAspHisMetArgCysAlaGlyAspValCysPheSerGluVal 140
DB 361 TCGTGGCAGGACCTTAAGGATCACATGCGCAAGCTGGAGATGCTGCTTCTTGAAGTT 420
QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLys 160
DB 421 TTCCCTGACCGTAAGGATGCTCTGGGTGTTGGATTATAGCACTATGATATGAAG 480
QY 161 TyrAlaileArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
DB 481 TACGCAATAAGGAACCTTGATGCCACTGAATTCGAAATGCTTCTCTAGTGTATATA 540
QY 181 ArgValArgGluTyrGlnSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
DB 541 CGGTGAGGGAATATGAGTCGAGGAGTGTGAGTGAAGCCAGATGATTTCTTAAAGCTAT 600
QY 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerVal 220
DB 601 AGAAGCAGAGTCGAGGCGGTGTCGAAGCTGTAGCTATAGTAGCAAGCAGAGGTG 660
QY 221 SerProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeuTyr 240
DB 661 TCACCTGTAGATCCATTTCCTCCGCTTCACGCGCCCTTAGTCGTTCTCGCTATAC 720
QY 241 SerSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTyrIle 256
DB 721 AGCTCTGTCTCAAGATCTGGCTCACTGCTACGAGCTGGGGATGGATC 768
RESULT 4
ID ABZ13394 standard; DNA; 762 BP.
XX AC ABZ13394;
XX DT 21-JAN-2003 (first entry)
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1199.
XX DE Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX KW Arabidopsis thaliana.
XX OS Arabidopsis thaliana.
XX PN WO200216655-A2.
XX PD 28-FEB-2002.
XX PF 24-AUG-2001; 2001WO-05026685.
XX PR 24-AUG-2000; 2000US-0227866P.
XX PR 26-JUN-2001; 2001US-0264647P.
XX PR 22-JUN-2001; 2001US-0300111P.
XX PA (SCRI) SCRIPPS RES INST.
XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX PI Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.
XX Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX Claim 144; SEQ ID NO 1199; 577pp + Sequence Listing; English.
XX The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX SQ Sequence 762 BP; 189 A; 161 C; 197 G; 215 T; 0 U; 0 Other;
Alignment Scores: 9.84e-88 Length: 762
Pred. No.: 1272.00 Matches: 253
Score: 91.67% Conservative: 0
Percent Similarity: 91.67% Mismatches: 1
Best Local Similarity: 91.67% Indels: 23
Query Match: 87.66% Gaps: 6
DB: 6
US-10-014-927-19 (1-279) x ABZ13394 (1-762)
QY 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyValAsnLeuProGlyAspIleArgLys 20
DB 1 ATCAGTAGCCGATGGATCGTACGATCTAGTTGGCAATTTGCTCGACATATTCGCAAG 60
QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspLeuAspLeuIle 40
DB 61 TGTGAGTTGAAGATCTCTTACAAAGTATGGACCAATTTGGACATTTGATTGAAGATT 120
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
DB 121 CCACCGAGACCTCTCGTTATGCGCTTTGTCAGATTGAAGATCTCTGATGACAGCAT 180
QY 61 AlaileTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
DB 181 GCATTTATGACGTGATGTTATGATTTGATGGGTGTCACCTTCGGTTGAGATTGCA 240
QY 81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSer 100
DB 241 CATGTGCTCGTAGATTTTACCATCATGTTAGTACAGTACAGCAGCAGTACAGTGCAGC 300
QY 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
DB 301 CGTGACCTTCAAGACGCTCTGACTACCGCTGCTGTGTGACCGGATTCACGCTTCGCT 360
QY 121 SerTyrGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
DB 361 TCGTGGCAGGACCTTAAGGATCACATGCGCAAGCTGGAGATGCTGCTTCTTGAAGTT 420
QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLys 160
DB 421 TTCCCTGACCGTAAGGATGCTCTGGGTGTTGGATTATAGCACTATGATATGAAG 480
QY 161 TyrAlaileArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
DB 481 TACGCAATAAGGAACCTTGATGCCACTGAATTTGAAATGCTTCTCTAGTGTATATA 540
QY 181 ArgValArgGluTyrGlnSerArgSerValSerArgSerProAspAspSerLysSerTyr 200
DB 541 CGGTGAGGGAATATGAGTCGAGGAGTGTGAGTGAAGCCAGATGATTTCTTAAAGCTAT 600

Qy 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerIysSerArgSerVal 220
Db 601 AGAAGCAGGAGTCGG
Qy 221 SerProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerIleuTyr 240
Db 627 TCACCTGCTAGATCCATTCCCGCGTTCACGGCCCTTAGTCGTTCTCGCTCGCTATAC 686
Qy 241 SerSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArg 260
Db 687 AGCTCTGTCTCAGG-----TCCCAATCAAGA 713
Qy 261 SerIysSerArgSerArgSerArgSerArgSerProValSerProVal 276
Db 714 TCAAAATCAAGATCAAGATCAAGATCGAATTCCTCCAGTTTCACCTGTA 761

RESULT 5
AAC42865
ID AAC42865 standard; DNA; 714 BP.
XX AAC42865;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 37144.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
FN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.

PF 20-APR-2000; 2000WO-AT000100.
 XX
 PR 23-APR-1999; 99AT-00000727.
 XX
 PA (OSTP) OESTERR FORSCH SEIBERSDORF.
 XX
 PI Barta A, Lopato S, Kalyana M, Dorner S;
 XX
 DR WPI; 2000-687349/67.
 XX
 DR P-PSDB; AAB11414, AAB11419.
 XX
 PT Novel proteins with splice-factor activity in plants, useful e.g. for
 PT altering flowering time or development, and the nucleic acid that encodes
 PT it.
 XX
 XX Claim 5; Fig 1a; 67pp; German.
 PS
 CC This invention describes a novel protein (I) with splice-factor activity
 CC in plants (I) modifies the choice of splice sites in many plant pre-
 CC mRNAs. (I) (also the nucleic acid that encodes them and related vectors
 CC or expression systems) are used; (i) to alter splice patterns in plants,
 CC or their parts; (ii) to alter developmental behavior of plants; and/or
 CC (iii) to delay flowering, particularly by at least 25% relative to the
 CC wild type, especially in crop plants such as cereals, beans, rice and
 CC fruit
 XX
 SQ Sequence 4044 BP; 1061 A; 728 C; 828 G; 1427 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 9,82e-57 Length: 4044
 Score: 875.00 Matches: 242
 Percent Similarity: 43.53% Conservative: 0
 Best Local Similarity: 43.53% Mismatches: 3
 Query Match: 60.30% Indels: 314
 DB: 3 Gaps: 9

US-10-014-927-19 (1-279) x AAC81899 (1-4044)

Qy	1	MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys	20
Db	808	ATGAGTAGCGAGTGAATCGATACGATACGTTGGGAATTTGCCGTGGAGATATTCGAAG	867
Qy	21	CysGluValGluAspLeuPheTyrLys-	29
Db	868	TGTGAGGTTGAAGATCTCTTACAGGTTTGAAATTTCCCTTTCTCTCGATAAAA	927
Qy	29	-	29
Db	928	TTGAAATTCATTATGACTAGTTTGGTTTCATAAATTTGCAATTCGTCTTCTCGACAAAT	987
Qy	30	-	39
Db	988	TTAAATCGACTCTTATGTATATTTGTTTCAGTATGCGACCAATTTGGACATTTGATTGAA	1047
Qy	39	SileProProArgProGlyTyrAlaPheValGlu-	51
Db	1048	GAITCCACCGAGACCTCTCGTTATGCTTTGTGCGA-GGTATATTGATCAAGTACAAAT	1106
Qy	51	-	51
Db	1107	TGTTTTTTCTTCTTCTTGTAATAGTATAGCTAATGACTAAGATGTTTGTATTGGTG	1166
Qy	52	-	70
Db	1167	GCAGTTTGAAGATCCTCGTGATGCGACGATGCAATTTATGGACGTGATGGTTATGATTT	1226
Qy	70	eAspGlyCysArgLeuArg-	76
Db	1227	TGATGGGTGTCGACTCTCGGTTAGTAAACCGCATGATGAAAGCTAGCTTAATTTCTGTGAA	1286
Qy	77	-	84
Db	1287	TTCTTGTAAGGTTTANCTTTGTGTGATGTTTTTTAGTTTGGATTGACATGCGTGGTC	1346

Qy	84	rgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSerArgAlaProS	104
Db	1347	GTAGATTTTCCACCATCAGTTGATAGGTACAGCAGCAGCTACAGTGGAGCGCGTGCACCTT	1406
Qy	104	erArgArgSerAspTyrArg-	110
Db	1407	CAGACGCTCTGACTACCGCGGTTTGTAGAGTCTTCTCGATTGTGTATTGTGTGTG	1466
Qy	111	-	115
Db	1467	TAAATTTTATATTGAAACTCATTTTACTACTAAACATGATGCTGTGTGACCGGA	1526
Qy	116	LeuProProSerAlaSerTrpGlnAspLeuLys-	126
Db	1527	TTACCGCTTCTGCTTGTGTCAGGACCTTAA-GGTAAAGGACACATATATAGTCTTTTTC	1585
Qy	126	-	126
Db	1586	TCTGAATGTTGGTCTCTATATATCATGTTTGGATTATCTCTTTCTGAAATGATGTTAT	1645
Qy	127	-	139
Db	1646	TTGCTATTACGGGTGATTAGGATCACATGCGCAAGCTGGAGATGCTCTGCTTCTGAA	1705
Qy	140	ValPheProAspArg-Lys-	145
Db	1706	GTTTTTCCCTGACCGTAAAGGTGAGTTGCATTCGATAGTTTGGATAAGCTTTTGTGATGA	1765
Qy	145	-	145
Db	1766	TGTGTTAGTAAATVAGTCTTTTGTGAAGGAGAAATAGGTGTTAAAGCATCTGAACCTAAAC	1825
Qy	146	-	158
Db	1826	TCACATTCAGTATTCTTTGTAGGCAATGCTGGGGTTGTGGATTATAGCAACTATGATGA	1885
Qy	158	pMetLysTyrAla-	162
Db	1886	TATGAAGTACGC-AGTAAAGTTTATATCTTTGCAACGCAAAATGTTCTCGACCTATGCT	1944
Qy	162	-	162
Db	1945	TAGACTGCTTTTGTTCATAGTATACGAGCTGAATTTATCTCTCGAGGCCAGTGTG	2004
Qy	163	-	170
Db	2005	GATCTTTGATGTTCCCTTAAATTTTGTGATGTGACAGATAAGGAAACTTGTATGCCACTGA	2064
Qy	170	upPheArgAsnAlaPheSerSerAlaTyrIleArg-	181
Db	2065	ATTTCGAAATGCTTTCTCTAGTGTATATACGGGTATGTTGTATTGCTTTCTTTGATTT	2124
Qy	182	-	186
Db	2125	TGTTAAGCATAAAGTGATATGAGTCTCTCTGAATTTACTGTTTCAGGTAGGGAATATG	2184
Qy	186	luSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgS	206
Db	2185	AGTCGAGAGTGTGAGTCCAAAGCCAGATGATTTCTAAAGCTATAGAACGAGGAGTCCGA	2244
Qy	206	erArgGlyProSerCysSerTyrSerSerLysSerArg-	218
Db	2245	GCCGTGGTCCAAAGCTGTAGCTATAGTACCAAGACGAGGAGTCTTTTTTTTTTTTTTT	2304
Qy	218	-	218
Db	2305	TTCTATAAACCTAAGACATATAAGGAGTTTATTATTGTAACCTTATTATGAAATTAACCTGAC	2364
Qy	219	-	233
Db	2365	TTCTAAATGCAATGACGTGTGTACCTGTAGATCATTTCCCGCGTTCACGGCCCTT	2424

QY 234 SerArgSerArgSerLeuTyrSerSerValSerArg 245
AAC34814
ID AAC34814 standard; DNA; 1232 BP.
XX
AC AAC34814;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 7984.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130691P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.

PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 23-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 12-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.

[illegible]

xx	7		Hybridisation assay; genetic mapping; gene expression control;
KW		Gaps:	KW protein identification; signal transduction pathway; metabolic pathway;
KW		(1-1232)	KW promoter; termination sequence; ss.
US-10-014-927-19	3	x AAC34814	

Qy 1 MetSerArgTrpAsnArgThrIleTyrValGlyAsnIleuProGlyAspIleArgLys 20
...

XX	EP1033405-A2.	PR	01-JUL-1999;	99US-0141842P.
PN		PR	01-JUL-1999;	99US-0142154P.
XX		PR	02-JUL-1999;	99US-0142055P.
PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390P.
XX		PR	08-JUL-1999;	99US-0142803P.
PF		PR	09-JUL-1999;	99US-0142920P.
PF		PR	12-JUL-1999;	99US-0142977P.
XX	25-FEB-2000; 2000EP-00301439.	PR	13-JUL-1999;	99US-0143542P.
XX		PR	14-JUL-1999;	99US-0143624P.
PR	25-FEB-1999;	PR	15-JUL-1999;	99US-0144005P.
PR	99US-0121825P.	PR	16-JUL-1999;	99US-0144085P.
PR	99US-012180P.	PR	16-JUL-1999;	99US-0144086P.
PR	99US-0123548P.	PR	19-JUL-1999;	99US-0144325P.
PR	99US-0125788P.	PR	19-JUL-1999;	99US-0144331P.
PR	99US-0126264P.	PR	19-JUL-1999;	99US-0144332P.
PR	99US-0126785P.	PR	19-JUL-1999;	99US-0144333P.
PR	99US-0127462P.	PR	19-JUL-1999;	99US-0144334P.
PR	99US-0128234P.	PR	19-JUL-1999;	99US-0144335P.
PR	99US-0128714P.	PR	20-JUL-1999;	99US-0144352P.
PR	99US-0130077P.	PR	20-JUL-1999;	99US-0144632P.
PR	99US-0130449P.	PR	20-JUL-1999;	99US-0144884P.
PR	99US-0130510P.	PR	20-JUL-1999;	99US-0144814P.
PR	99US-0130891P.	PR	21-JUL-1999;	99US-0145086P.
PR	99US-0131449P.	PR	21-JUL-1999;	99US-0145088P.
PR	99US-0132048P.	PR	21-JUL-1999;	99US-0145085P.
PR	99US-0132407P.	PR	22-JUL-1999;	99US-0145087P.
PR	99US-0132484P.	PR	22-JUL-1999;	99US-0145089P.
PR	99US-0132485P.	PR	22-JUL-1999;	99US-0145192P.
PR	99US-0132486P.	PR	22-JUL-1999;	99US-0145145P.
PR	99US-0132487P.	PR	23-JUL-1999;	99US-0145218P.
PR	99US-0132863P.	PR	23-JUL-1999;	99US-0145224P.
PR	99US-0134258P.	PR	23-JUL-1999;	99US-0145276P.
PR	99US-0134218P.	PR	26-JUL-1999;	99US-0145913P.
PR	99US-0134221P.	PR	27-JUL-1999;	99US-0145918P.
PR	99US-0134370P.	PR	27-JUL-1999;	99US-0145919P.
PR	99US-0134768P.	PR	28-JUL-1999;	99US-0145951P.
PR	99US-0134941P.	PR	02-AUG-1999;	99US-0146386P.
PR	99US-0135124P.	PR	02-AUG-1999;	99US-0146388P.
PR	99US-0135353P.	PR	02-AUG-1999;	99US-0146389P.
PR	99US-0135629P.	PR	03-AUG-1999;	99US-0147038P.
PR	99US-0136021P.	PR	04-AUG-1999;	99US-0147204P.
PR	99US-0136392P.	PR	04-AUG-1999;	99US-0147202P.
PR	99US-0136782P.	PR	05-AUG-1999;	99US-0147192P.
PR	99US-0137222P.	PR	05-AUG-1999;	99US-0147260P.
PR	99US-0137528P.	PR	06-AUG-1999;	99US-0147303P.
PR	99US-0137502P.	PR	06-AUG-1999;	99US-0147416P.
PR	99US-0137724P.	PR	09-AUG-1999;	99US-0147493P.
PR	99US-0138094P.	PR	09-AUG-1999;	99US-0147935P.
PR	99US-0138540P.	PR	10-AUG-1999;	99US-0148171P.
PR	99US-0138847P.	PR	11-AUG-1999;	99US-0148319P.
PR	99US-0139119P.	PR	12-AUG-1999;	99US-0148341P.
PR	99US-0139452P.	PR	13-AUG-1999;	99US-0148565P.
PR	99US-0139453P.	PR	13-AUG-1999;	99US-0148684P.
PR	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	99US-0139817P.	PR	30-AUG-1999;	99US-0151080P.
PR	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	99US-0139899P.	PR	31-AUG-1999;	99US-0151438P.
PR	99US-0140353P.	PR	01-SEP-1999;	99US-0151930P.
PR	99US-0140354P.	PR	07-SEP-1999;	99US-0152363P.
PR	99US-0140695P.	PR	10-SEP-1999;	99US-0153070P.
PR	99US-0140823P.	PR	13-SEP-1999;	99US-0153758P.
PR	99US-0140991P.	PR	15-SEP-1999;	99US-0154018P.
PR	99US-0141287P.			

PR	23-APR-1999;	99US-0130510P.	PR	20-JUL-1999;	99US-0144632P.
PR	23-APR-1999;	99US-0130891P.	PR	20-JUL-1999;	99US-0144884P.
PR	28-APR-1999;	99US-0131449P.	PR	21-JUL-1999;	99US-0144814P.
PR	30-APR-1999;	99US-0132048P.	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	99US-0132407P.	PR	22-JUL-1999;	99US-0145088P.
PR	04-MAY-1999;	99US-0132484P.	PR	22-JUL-1999;	99US-0145085P.
PR	05-MAY-1999;	99US-0132485P.	PR	22-JUL-1999;	99US-0145087P.
PR	06-MAY-1999;	99US-0132486P.	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999;	99US-0132487P.	PR	22-JUL-1999;	99US-0145192P.
PR	07-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145145P.
PR	11-MAY-1999;	99US-0134256P.	PR	23-JUL-1999;	99US-0145218P.
PR	14-MAY-1999;	99US-0134219P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134221P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145918P.
PR	18-MAY-1999;	99US-0134768P.	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	99US-0134941P.	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	99US-0135629P.	PR	03-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	99US-0136021P.	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	99US-0136393P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0136782P.	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139119P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.

```
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 9,62e-55 Length: 885
Score: 838.50 Matches: 180
Percent Similarity: 72.07% Conservative: 29
Best Local Similarity: 62.07% Mismatches: 43
Query Match: 57.79% Indels: 39
DB: 3 Gaps: 7

US-10-014-927-19 (1-279) x AAC42654 (1-885)

Qy 1 MetSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
Db 1 ATGAGCAGCGTTCAGTAGAAGATTACGTGCGGAACCTTCCCGCGATATCCGTGAA 60
Qy 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLysLys 40
Db 61 AGAGAAGTTGAAGACTTGTTCAGTAAGTATGGAACCTGTTGTTCAAAATCGATTGAAGATT 120
Qy 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60
Db 121 CCGCGAGGCTCCAGGCTATGCACTGCTGAGTTGAGGATGCTGCTGATGCTGATGAT 180
Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db 181 GCAATTTATGCGCGTATGTTATGATGATGATGATGATGATGATGATGATGATGAT 240
Qy 81 HisGlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAla--- 99
Db 241 CATGCTGGAGCGCTTCATCATCATGATGCA-----CGCGTAGTTATAGTGTGCT 291
Qy 100 -----SerArgAlaProSerArgArg 106
Db 292 GGTCTGGCGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 351
Qy 107 SerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLys 126
Db 352 TCAGATACCCGTTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGAT 411
Qy 127 AspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGly 146
Db 412 GATCAGATCGCTAAAGGAGGAGAGTGTGTTTCTCAAGTGTGTTGATGATGATGAT 471
Qy 147 MetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArg----- 164
Db 472 ACAACTGGAAATGTAGATTATACAGGATACAGGACATGAATATGCGGTGAG-ATATAA 530
Qy 165 ---LysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArg 183
Db 531 AAAAAGCTCGATGACACAGAGTTTTCGGAATGCGTTTCTCATGATATGTTCCGGTTAGA 590

184 GluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyr---ArgSer 202
591 GAATATGATTACAGAGGAGGATTTCGAGGAGCCCGCGTGAAGATCCTATTCTTAAGAGC 650
203 ArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSerPro 222
651 CGCAGCCGTGAGGAGTCCAGC-----CGTAGTCGTAGCGCAGCAGAGCAGG 701
223 AlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSer 242
702 AGCAAGAGCAGAGTCCAAAGGCTAAATCTTTGGTAGATCGCTGCA-----AAATCT 755
243 ValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSerLys 262
756 ACATCGAGATCT-----CCTCGCTCTCGC 779

263 SerArgSerArgSerArgSerAsnSerPro 272
780 TCCGCTCTAAGTCGAGGTCACTGTCTCCA 809

RESULT 10
AAC37026
ID AAC37026 standard; DNA; 1241 BP.
XX AC AAC37026;
XX AC
DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15916.
XX KW Hybridisation assay; genetic mapping; gene expression control;
XX KW protein identification; signal transduction pathway; metabolic pathway;
XX KW promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 200EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 28-APR-1999; 99US-0130891P.
XX PR 30-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 04-MAY-1999; 99US-0132407P.
XX PR 05-MAY-1999; 99US-0132484P.
XX PR 06-MAY-1999; 99US-0132485P.
XX PR 06-MAY-1999; 99US-0132488P.
XX PR 06-MAY-1999; 99US-0132487P.
XX PR 07-MAY-1999; 99US-0132863P.
XX PR 11-MAY-1999; 99US-0134256P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134221P.
XX PR 14-MAY-1999; 99US-0134370P.
XX PR 18-MAY-1999; 99US-0134768P.
XX PR 19-MAY-1999; 99US-0134941P.
XX PR 20-MAY-1999; 99US-0135124P.
XX PR 21-MAY-1999; 99US-0135353P.
```

PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142300P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.

PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 23-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 13-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 23-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.

PR 28-OCT-1999; 99US-01611992P.
PR 28-OCT-1999; 99US-01611993P.
PR 29-OCT-1999; 99US-01621422P.

Alignment Scores:
Pred. No.: 5,01e-52 Length: 1241
Score: 805.00 Matches: 183
Percent Similarity: 71.67% Conservative: 27
Best Local Similarity: 62.46% Mismatches: 53
Query Match: 55.48% Indels: 30
DB: 10 Gaps: 10

US-10-014-927-19 (1-279) x AAC37026 (1-1241)

QY 1 MetSerArgTrpAenArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
DB 141 ATGAGTGGCGGATTTCTCGGTCAATCTATGTTGGTAACCTGCCCGGTGACATTAGGGA 200

QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLys 40
DB 201 CATGAGATTCAAGATATCTTTTACAAGTATGGCCGCAATGTCGATATTGAATTGAAGGT 260

QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60
DB 261 CCACCTCGGCTCCATGTTATTCGTTGTTGAGTTTGAGCATCTCGGATGCTGAAGAT 320

QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluLeuAla 80
DB 321 GCCATCAAGGCCGTGATGCTATTAATTCGATGGCTGCTGCTGAGGTTGAGCTTGA 380

QY 81 HisGlyGlyArgArgPheSerProSerValAspArgTyr-----Ser 94
DB 381 CATGGTGGTCGAGGACAGTCTTCAAGT---GATCGTGGTGGTACGCTGCTGGTGC 437

QY 95 SerSerTyr-----SerAlaSerArgAlaProSerArgArg 106
DB 438 AGCGGTATGGTGGTGGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 497

QY 107 SerAspTyrArgValLeuValThrGlyLeuProSerAlaSerTrpGlnAspLeuLys 126
DB 498 TCTGAATTCGAGGATATTGTACGTTGGCTCCCATCATCTGCTCATGGCAAGATTGA 557

QY 127 AspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGly 146
DB 558 GATCATATGCGAAGCTGTGATGTGTTGCTGAGGTGACTCGACACATGATGGA 617

QY 147 MetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeu 166
DB 618 ACTTATGGTGTTCGACTACCACTTATGATGACATGAATGATGCAATAGGAAACTT 677

QY 167 AspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGlu 186
DB 678 GATGACACAGAGTTTCAGAAACCCCTGGGCTAGAGGTTTATCCGGGTTAAGAAATATGA 737

QY 187 ---SerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArg 205
DB 738 AGCTCCCGATCAAGAGCAGAGCCCAAGCAGAGAGCAAGT---CGAATGCTGATGCCGA 794

QY 206 SerArgGlyProSerCysSerTyrSer---SerLysSerArgSerValSerProAlaArg 224
DB 795 GGCGGTGTTCGACCCATAGCCGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 854

QY 225 SerIle-----SerProArg-----SerArgProLeuSerArgSerArgSerLeu 239
DB 855 GATCTGAGTAAATCACCAGGCGATCCCTTCCAGATCGATTTCCAAATCTAGATCG--- 911

QY 240 TyrSerSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSer 259
DB 912 ---CCATCGCCGCAAGAGAGAGTCCCGCCAGGGCA-----ATGTCG 953

QY 260 ArgSerLysSerArgSerArgSerArgSerArgSerArgSerPro 272
DB 954 AGATCAAAAGTCCAGGTCGAGTCCAGGTCGAGTCTCGG 992

RESULT 11
ADD22452
ID ADD22452 standard; DNA; 2765 BP.
XX AC ADD22452;
XX DT 15-JAN-2004 (first entry)
XX DE HLA-B46 T cell recognised tumour antigenic polypeptide, SEQ No 102.
XX KW tumour antigenic peptide; cancer; vaccine; cytostatic; cytotoxic T cell;
XX KW colon; mouth; lung; prostatic; gynecological; human; gene; ds.
XX OS Homo sapiens.
XX PN JP2003111595-A.
XX PD 15-APR-2003.
XX PF 24-JUN-2002; 2002JP-00183603.
XX PR 25-JUN-2001; 2001JP-00191974.
XX PA (ITOY/) ITO Y.
XX WI; 2003-611129/58.
XX PT Novel tumor antigenic peptide or polypeptide useful for inducing
PT cytotoxic T cells or for treating cancer such as colon, mouth, lung,
PT prostatic or gynecological cancer.
XX PS Claim 10; SEQ ID NO 102; 98pp; Japanese.
XX CC The invention relates to a novel tumour antigenic peptide or polypeptide
CC comprising a sequence selected from 99 sequences fully defined in the
CC specification. The tumour antigenic peptide or polypeptide comprises a
CC sequence selected from 99 sequences fully defined in the specification,
CC where the tumour antigenic peptide preferably has a sequence of Glu-Pro-
CC Pro-Leu-Ser-Gln-Glu-Thr-Phe, and the polypeptide preferably has a
CC sequence comprising 393 amino acids fully defined in the specification.
CC The invention further provides a cancer vaccine comprising a tumour
CC antigenic peptide or polypeptide, which has cytostatic activity. A tumour
CC antigenic peptide, polypeptide, its encoding polynucleotide, a
CC hybridising polynucleotide, a recombinant vector containing the
CC polynucleotide, a host transformed with the vector or an antibody are
CC useful for screening for compounds that interact with the tumour
CC antigenic peptide, the polypeptide or its encoding polynucleotide and
CC increases the expression of the tumour antigenic peptide, the polypeptide
CC or polynucleotide. The tumour antigenic peptide or the polypeptide is
CC useful for inducing cytotoxic T cells. The tumour antigenic peptide
CC vaccine is useful for treating cancer such as colon, mouth, lung,
CC prostatic or gynecological cancer. The invention also provides a
CC pharmaceutical composition useful for treating cancer. The tumour
CC antigenic peptide or the polypeptide is useful as an antigen to create
CC antibodies. This polynucleotide sequence represents the DNA encoding one
CC of the human tumour antigenic polypeptides of the invention.

XX SQ Sequence 2765 BP; 759 A; 500 C; 652 G; 854 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 7,13e-35 Length: 2765
Score: 584.50 Matches: 134
Percent Similarity: 69.07% Conservative: 29
Best Local Similarity: 56.78% Mismatches: 60
Query Match: 40.28% Indels: 13
DB: 9 Gaps: 7

US-10-014-927-19 (1-279) x ADD22452 (1-2765)

QY 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
DB 142 ATCTAGTGGTAACTTACCTCCAGACATCCGACCAAGGACATTGAGGACGCTGTCTAC 201

[illegible]

Qy	9	IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr	28
Db	176	ATCTACGTGGTAACTTACTCTCCAGACATCCGAACCAAGACATTGAGGACGTGTCTTAC	235
Qy	29	LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla	48
Db	236	AAATACGGCGCTATCCGACATCACTCAAGATCCCGCGGGGACCGCCCTTCGCC	295
Qy	49	PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr	68
Db	296	TTCTGTTGAGTTCAGGACCCGAGACGGGAAGACGGCGGTATGTGTCGACGCGCTAT	355
Qy	69	AspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGlyArg	84
Db	356	GATTACGATGGGTACCGTCTCGGGTGGAGTTTCCTCGAAGCGCGCGTGGAAACAGCCGA	415
Qy	85	-----ArgPheSerProSerValAsp	91
Db	416	GGCGCGCGGGGTGGAGTGGCGGAGTCCCGGAGTTCGTATGGCCCC-----	466
Qy	92	ArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgArgSerAspTyrArgVal	111
Db	467	-----CCATCCAGCGCGTCTGAAAACAGAGTG	493
Qy	112	LeuValThrGlyLeuProProSerAlaSerTyrGlnAspLeuLysAspHisMetArgLys	131
Db	494	GTTCGTCTCTGGACTGCTCCAAAGTGGAAAGTTGGCAGATTAAAGGATCACATGGGTGAA	553
Qy	132	AlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyValVal	151
Db	554	GCAGGTGATGTATGTTATGCTGATGTTTACCGAGATGGC-----ACTGGTGTCTGTG	604
Qy	152	AspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPhe	171
Db	605	GAGTTTCTACGGAAGAGATATGACATATGACATTGAGTTTCAAAAACATGATATAACATAAGTTT	664
Qy	172	Arg---AsnAlaPheSerSerAlaTyrIleArgValArgGluTyrGluSerArgSerVal	190
Db	665	AGATCTCATGAGGAGAAACATGCTCATCCGGGTTAAAGTTGATGGGCCCAAGTCCA	724
Qy	191	SerArgSerProAspAspSerLysSerTyrArgSerArgSerArgGlyProSer	210
Db	725	AGTTATGGAAAGATCTCGATCTCGAAGC---CGTAGTCGTAGCAAGCCGTACGAGAAGC	781
Qy	211	CysSerTyrSerSerLysSerArgSerValSerProAlaArgSerIle---SerProArg	229
Db	782	-----AACAGCAGGAGTCCAGTTACTCCCCAAGGAGAGACAGAGATCACCACGC	832
Qy	230	SerArgProLeu---SerArgSerArgSer	238
Db	833	TATTCTCCCGTCATAGCAGATCTCGCTCT	862
RESULT 14			
ABL29431	standard; DNA; 1478 BP.		
XX	AC	ABL29431;	
XX	XX		
DT	26-MAR-2002	(first entry)	
XX	XX		
DE	XX	Drosophila melanogaster genomic polynucleotide SEQ ID NO 39766.	
XX	XX	Drosophila; developmental biology; cell signalling; insecticide;	
KW	KW	pharmaceutical; gene; ds.	
XX	XX	Drosophila melanogaster.	
OS	XX		
XX	XX	WO200171042-A2.	
XX	XX		
PD	27-SEP-2001.		
XX	XX		
FF	23-MAR-2001; 2001WO-US009231.		

Db	540	GTGTAGTTCTTCGGCCAGAGGACATCAAGTACGC	CAATCAAAAAATTGACGACTCTCCG	599
Qy	171	PheArgAsnAlaPheSerSer--AlaTyrIleArgValArgGluTyrGluSerArgSer	189	
Db	600	TTCCCATCGCATGAGGCGGAGGTTCCTACATTCGGTACGCAG--GATACGGTGAT	656	
Qy	190	ValSerArg-----SerProAsp	195	
Db	657	AACGACAGAGCGGTGCTGGCGCGGACGGTGGAGGTGTCGGCGGACGGCGGAGGT	716	
Qy	196	AspSerLysSerTyrArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSer	215	
Db	717	GGAAGCGGTGACTACCGGACAGATCCCGTTCGGC-----TCATTCG	761	
Qy	216	LysSerArgSerValSerProAlaArgSerIleSerProArgSerArgProLeuSerArg	235	
Db	762	-----TCGAGGCGCGCGCGT	779	
Qy	236	SerArgSerLeuTyrSerSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrp	255	
Db	780	GGCACTCCACACTACTCCCGACATCGCAGT-----	809	
Qy	256	IleSerGlnSerArgSerLysSerArgSerArgSerArgSerArgSerProValSerPro	275	
Db	810	-----CAATCTTTATCCAGGTCTCGCTCACTTCACTTAATATAACACAACCA	854	
Qy	276	ValIle	277	
Db	858	AGCATT	863	

RESULT 15
ABI99368
ID ABI99368 standard; CDNA; 1584 BP.
XX
AC ABI99368;
XX
XX
07-MAR-2002 (first entry)
XX

DE	Mouse ischaemic condition related cDNA sequence SEQ ID NO:1264.
XX	
XX	Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KW	vasospastic ischaemia; ischaemic condition; ischaemic disease; es.
XX	
XX	
OS	Mus musculus.
XX	
XX	WO20018188-A2.
PN	
XX	
PD	22-NOV-2001.
XX	
XX	
PF	18-MAY-2001; 2001WO-JP004192.
XX	
XX	
PR	18-MAY-2000; 2000JP-00145977.
XX	
XX	(UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
PA	
XX	Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
PI	
XX	WPI; 2002-034733/04.
DR	
XX	
PT	Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT	expression levels of particular genes defined in the specification or by
PT	determining the expression profile of a gene group comprising these
PT	genes.
XX	
PS	Claim 2; Page 745; 2690pp; English.
XX	
XX	
CC	The present invention describes a method for examining ischaemic
CC	conditions, comprising measuring the expression levels of particular
CC	genes (i) in a test sample or determining the expression profile of a
CC	gene group in the sample comprising genes selected from (i). The method
CC	is useful for examining the ischaemic condition (e.g. compressive
CC	ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC	expression levels of particular genes (ABI99202 to ABI99912 encoding the

CC protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The expression
CC levels or expression profiles produced by these genes are used as an
CC indicator when screening for ischaemic condition-improving drugs or
CC therapeutics for ischaemic diseases. ABI99913 and ABI99914 represent PCR
CC primers for a mouse ischaemic condition related sequence, which are used
CC in the exemplification of the present invention

XX SQ Sequence 1584 BP; 374 A; 332 C; 428 G; 450 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5,22e-32 Length: 1584
Score: 543.00 Matches: 139
Percent Similarity: 52.68% Conservative: 28
Best Local Similarity: 43.85% Mismatches: 42
Query Match: 37.42% Indels: 109
DB: 6 Gaps: 9

US-10-014-927-19 (1-279) x ABI99368 (1-1584)

QY 9 IletyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
DB 131 ATCTACGTGGTAACTTACCTCCGGATATCGAACCAAGACATCGAGGACGTGTTTAC 190
QY 29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArg-ProProGlyTyrAl 48
DB 191 AAATACGGCGCATCCGACATCGACTGAAGAA-CGGCGGGGGGACCGCCCTCGC 249
QY 48 aPheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyr 68
DB 250 CTTTCGTTGAGTTCGAGGACCGCGAGACCGGAGAGATCGCGTGTACGTCGCGACGGCTA 309
QY 68 rAspPheAspGlyCysArgLeuArgValGluAlaHisGlyGlyArg----- 84
DB 310 CGACTACGACGGCTACCGGCTGCGGGTAGAGTTTCCCGAAGCGCGCGGACCGGCGCG 369
QY 85 -----ArgPheSerProSerValAs 91
DB 370 AGCGCGCGGGGGTGGAGCGCGCGCGCGCGCGCGCTATGCGCGG----- 421
QY 91 pArgTyrSerSerTyrSerAlaSerAlaProSerArgArgSerAspTyrArgVa 111
DB 422 -----CCGTCCAGCGGTCCGAGAACACAGT 447
QY 111 lLeuValThrGlyLeuProProSerAlaSerTyrGlnAspLeuLysAspHisMetArgLy 131
DB 448 GGTTCCTCTCGACTGCTCCGAGTGGAGGCTGCGAGACTTAAAGGATCACATGCGGTGA 507
QY 131 eAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyValVa 151
DB 508 GCGAGGTGATGTATGTACGCTGATGTTTACCGAGATGCG-----ACTGGTGTCGT 558
QY 151 lAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPh 171
DB 559 GGAGTTTGTACGGAAGAGATATGACGTATGCACTTCGAAACTGGATAACACTAAGTT 618
QY 171 eArg-----Asn-AlaPheSer- 176
DB 619 TAGATCTCAGGAGGTAGGTATACACTTATCTTTTTCGCGAGAAATGGATACAGTT 678
QY 176 ----- 176
DB 679 TTCTTAACAGTGGAAATTGAAGGTAAGGATACAGCAAGGTGTTTCGTAATATACCAG 738
QY 176 ----- 176
DB 739 AGCCCTGATCTGCTTTGTATTGCTTGTAGCTTGTCTGAGACAGGTGAAAGCTTAGATCT 798
QY 177 -----SerAlaTyrIleArgValArgGluT 185
DB 799 TTCAATGGAAAGTTCTGTCTATCCAAATAGGAGAAACTGCGCTACATCCGGGTTAAAGTTG 858
QY 185 yrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSera 205

DB 859 ATGGGCCCCAGAGTCCAGAGTTATGGAAGATCTCGATCTCGAAGC---CGTAGTCGTAGCA 915
QY 205 rgSerArgGlyProSerCysSerTyrSerLysSerArgSerValSerProAlaArgS 225
DB 916 GAGGCCGTAGCAGAAGC-----AACAGCAGGAGTCCGAGTTACTTCCCAAGAGAA 966
QY 225 erIle---SerProArgSerArgProLeu---SerArgSerArgSer 238
DB 967 GCAGAGGATCACCGCTATTCTCCCGTCATAGCAGATCTCGCTCT 1013

Search completed: March 26, 2004, 18:26:37
Job time : 505 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 26, 2004, 18:17:24 ; Search time 109 Seconds
(without alignments)
1420.471 Million cell updates/sec

Title: US-10-014-927-19

Perfect score: 1451
Sequence: 1 MSSRNRTTYVGNLPGDIRK.....RKSRSRSPVSPVWISG 279

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10014927/runat_24032004_152926_9928/app_query.fasta_1.455
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10014927 -CGEN_1_69 @runat_24032004_152926_9928 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DISPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/prodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/2/ina/PTUS.COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C	1	541.5	37.3	5285	4 US-09-402-328-1
	2	418.5	28.8	1534	4 US-09-976-594-961
	3	254.5	17.5	284	4 US-09-313-294A-5971
	4	214.5	14.8	2886	4 US-09-976-594-960
	5	208	14.3	955	4 US-09-620-312D-524
	6	187.5	12.9	193	4 US-09-702-705-1719
	7	187.5	12.9	193	4 US-09-736-457-1719
	8	187.5	12.9	193	4 US-09-671-325-1719
	9	179	12.3	1015	4 US-09-623-655-867
	10	165	11.4	479	4 US-09-621-976-3128
	11	160	11.0	1894	2 US-08-935-450-7
	12	160	11.0	1894	4 US-09-338-123-7

13	154.5	10.6	1278	3	US-09-613-182-14	Sequence 14, Appl	
14	154.5	10.6	2846	3	US-09-613-182-5	Sequence 5, Appl	
15	146.5	10.1	488	4	US-09-621-976-2159	Sequence 2159, Ap	
16	146	10.1	6775	4	US-09-620-312D-289	Sequence 289, App	
17	143.5	9.9	320	4	US-09-313-294A-3881	Sequence 3881, Ap	
18	139	9.6	3283	3	US-09-061-709-8	Sequence 8, Appl	
19	139	9.6	3283	4	US-09-899-651-8	Sequence 8, Appl	
20	139	9.6	3412	3	US-09-061-709-6	Sequence 6, Appl	
21	139	9.6	3412	4	US-09-899-651-6	Sequence 6, Appl	
22	137	9.4	2228	1	US-07-726-607C-1	Sequence 1, Appl	
23	137	9.4	2228	1	US-07-843-949A-1	Sequence 1, Appl	
24	137	9.4	2228	2	US-08-218-978-1	Sequence 1, Appl	
C	25	136.5	9.4	4248	4	US-10-164-595-53	Sequence 53, Appl
26	133	9.2	3034	4	US-09-976-594-636	Sequence 636, App	
27	132.5	9.1	1740	4	US-09-643-597-347	Sequence 347, App	
28	132.5	9.1	1740	4	US-09-542-615A-347	Sequence 347, App	
29	132.5	9.1	1740	4	US-09-606-421B-347	Sequence 347, App	
30	131.5	9.1	4159	3	US-09-061-709-4	Sequence 4, Appl	
31	131.5	9.1	4159	4	US-09-899-651-4	Sequence 4, Appl	
32	131.5	9.1	4181	4	US-09-643-597-175	Sequence 175, App	
33	131.5	9.1	4181	4	US-09-480-884A-175	Sequence 175, App	
34	131.5	9.1	4181	4	US-09-542-615A-175	Sequence 175, App	
35	131.5	9.1	4181	4	US-09-606-421B-175	Sequence 175, App	
36	130.5	9.0	781	4	US-09-370-838-150	Sequence 150, App	
37	130.5	9.0	1401	1	US-07-843-949A-3	Sequence 3, Appl	
38	130.5	9.0	1401	2	US-08-218-978-3	Sequence 3, Appl	
39	130.5	9.0	1477	4	US-09-976-594-794	Sequence 794, App	
40	128	8.8	522	4	US-09-621-976-3652	Sequence 3652, Ap	
41	127.5	8.8	547	4	US-09-370-838-168	Sequence 168, App	
42	127	8.8	582	3	US-09-385-982-421	Sequence 421, App	
43	126.5	8.7	548	4	US-09-370-838-174	Sequence 174, App	
44	126	8.7	1467	1	US-07-881-075-50	Sequence 50, Appl	
45	126	8.7	1467	1	US-08-120-827-50	Sequence 50, Appl	

ALIGNMENTS

RESULT 1

US-09-402-328-1/c
; Sequence 1, Application US/09402328
; Patent No. 6365728

GENERAL INFORMATION:

APPLICANT: Purdue Research Foundation,
Hodges, Thomas K.
Lysnik, Leszek A.

TITLE OF INVENTION: Regulatory Element For Expressing Genes
In Plants

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Barnes & Thornburg
STREET: 11 S. Meridian
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46204

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/402,328

FILING DATE: 05-No. 6365728-1999

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Breen, John P.

REGISTRATION NUMBER: 38, 833

REFERENCE/DOCKET NUMBER: 3220-29933

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317) 231-7745

TELEFAX: (317) 231-7433

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

; LENGTH: 5285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-402-328-1

Alignment Scores:
Pred. No.: 2,26e-39 Length: 5285
Score: 541.50 Matches: 180
Percent Similarity: 35.31% Conservative: 28
Best Local Similarity: 30.56% Mismatches: 44
Query Match: 37.32% Indels: 338
DB: 4 Gaps: 14

US-10-014-927-19 (1-279) x US-09-402-328-1 (1-5285)
Qy 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
Db 3693 ATGAGCAGCCGTTCCAGTAGAACGATTACGTCGGGACCTCCCGCGGATATCCGTGAA 3634
Qy 21 CysGluValGluAspLeu----- 26
Db 3633 AGAAGAGTTGAAGACTTGTTCAGTAAGGTAATAATCTCGTTCTCTCTAATTTAAGAACTC 3574
Qy 26 ----- 26
Db 3573 AGAACTCAGGATCGAATTTTGTGTTATGTGATGATATGTCCTTTACAAGATTTCGATGTG 3514
Qy 26 ----- 26
Db 3513 AAAGAAATGAATGATCTTCGGAATGTTGGATTCTGATTTTATTTGGAAGAAGTTTCC 3454
Qy 27 -----PheTyrIleSyrGlyProIleValAspIleAspLeu 38
Db 3453 TTAAGAAATGTTCTTCCTTATTTTGTG-AAGTATCGACCTGTGTTTCAAAATCGATTG 3395
Qy 39 LysIleProProArgProGlyTyrAlaPheValGlu----- 51
Db 3394 AAGATTCCGCGAGCCCTCCAGGCTATGCAATTCGTCGAGGTCAGATTTTATGCAACCA 3335
Qy 51 ----- 51
Db 3334 CATGTTTTATAAGSCITTTGTTGATTATTTACAGTTAGGCTCTTCGTTGTTGTAGTGT 3275
Qy 51 ----- 51
Db 3274 TCATGAGTTTCAAAATGTTGGAGATACCTACCATAATGTTTGTGATATATGATTGA 3215
Qy 52 -PheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAspPheAs 71
Db 3214 GTTTGAGGATGCTCGTGTGCTGATGATGCAATTTATGCGCGTATGATGATTGATGCTTGA 3155
Qy 71 pGlyCysArgLeuArg----- 76
Db 3154 TGGGCATCATTTACGGTGTGTTTAAATCATCAAAATTCAGAAATTTTACACAAACTTTGTA 3095
Qy 77 -----ValGluIleAlaH 81
Db 3094 TTGTTGACTGGTAGCAACATAAATTTATGTGTGTTGGCCATTGCGAGGTGGAAGTAGCTC 3035
Qy 81 isGlyGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerAla---- 99
Db 3034 ATGGTGGGAGCGGTTTCATCATCATGATGCA-----CGCGGTAGTTATAGTGGTGTG 2984
Qy 100 -----SerArgAlaProSerArgArg 107
Db 2983 GTTCGTGGCGGTGCTGGTGTGTGACGGTGTGTGTCGTCGAACGTGGACCACTTAGGAGAT 2924

Qy 107 erAspTyrArg----- 110
Db 2923 CAGAGTACCGCGGTACATATGATATGTTAGCTTTAGTATTATCATAGTTTATAGGAGAA 2864
Qy 111 -----ValLeuValThrGlyLeuProProSerAlas 121
Db 2863 ATCACTAAATAAGATTTTCCACTATACAGTTGTAGTGTGAGGTTTGCCTTCATCTCGT 2804
Qy 121 erTyrGlnAsp----- 124
Db 2803 CCGCAAGACCTCAAGGTGATTACACAGTTGTGGACAGATTCTCTAATTCGTTCTCTAT 2744
Qy 125 -----LeuLysAspHisMetArgLys 131
Db 2743 ATTTTCTACGATTTCTGAATGTTGACATATCACCTTTGTTTGAAGGATCATATCGGTAAA 2684
Qy 132 AlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMetSer----- 148
Db 2683 GGAGGAGAAGTTTGTCTTCAAGTGTTCGTGATGTTAGAGGTAAATCTCATGCTT 2624
Qy 148 ----- 148
Db 2623 ACCCAAGAAATAGAGATGTCTAAAGCTGTTTGGGTTTCTGCCTTTAAATCTCTGTTTA 2564
Qy 148 ----- 148
Db 2563 ACTAGTCTGTCATGCTTTAAAAAATCTATATCGTGGACTATTTATTATACAGGTACAAC 2504
Qy 149 GlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArg----- 164
Db 2503 GGAATTTAGATTATACCACTACAGGACATGAATATGCGGTGAGATGTTTCTTCTCC 2444
Qy 164 ----- 164
Db 2443 TCTTCATATTATTGATTTTCTGTGGGACTGTGGCTGAGAATGTGTTCATCTCTGATT 2384
Qy 165 -----LysLeuAspAlaThrGluPheArg 172
Db 2383 ACTCATGATTATATCTTTTATTTTGTCCAGATAAAAAAAGCTCGATGACACAGAGTTTCG 2324
Qy 172 glnAlaPheSerSerAlaTyrIleArg----- 181
Db 2323 GAATCGGTTTCTCATGAATATGTTGGGTATGTTAAAGACAAGAAATCTCGAGCTGCA 2264
Qy 182 -----ValArgGluTyrGluSerA 188
Db 2263 TTCTGTTTTATTGCACTTAACATTTCCCTATGTCTATTACAGTTAGAGAATATGATCAA 2204
Qy 188 rgSerValSerArgSerProAspAspSerLysSerTyr---ArgSerArgSerArgSerA 207
Db 2203 GGAGGGAATTCGAGGAGCCCGCCGCGTGAAGATCCTATTCTAAGAGCGCGACCGCTGAC 2144
Qy 207 rgGlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSerIleS 227
Db 2143 CGAGTCCCAAGC-----CGTAGTCGTAGCGCAGCAGAGAGGAGGAGGACAGAGCAA 2093
Qy 227 erProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerArgSerG 247
Db 2092 GTCCAAAGGCTAAATCTTTGCGTAGATCGCTGCA-----AAATCTACATCGATCT- 2040
Qy 247 LysSerLeuLeuArgAlaGlyAspTyrPheSerGlnSerArgSerLysSerArgSerArgS 267
Db 2039 -----CCTCGCTCTCGCTCCGCTCTAAGT 2015

RESULT 2
US-09-976-594-961
; Sequence 961, Application US/09976594
; Patent No. 6673549

GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: 09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 961
; LENGTH: 1534
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 988704.26
US-09-976-594-961

Alignment Scores:
Pred. No.: 6,33e-29 Length: 1534
Score: 418.50 Matches: 121
Percent Similarity: 52.63% Conservative: 29
Best Local Similarity: 42.46% Mismatches: 90
Query Match: 28.84% Indels: 45
DB: 4 Gaps: 10

US-10-014-927-19 (1-279) x US-09-976-594-961 (1-1534)
QY 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
DB 188 GTATTTCATCGGAGATTAATCCAGCGCCAGGAGAGAGCGTGGAAAGATTCTTCAAG 247
QY 29 LysTyrGlyProIleValAspIleAspLeuLysIleProArgProGlyTyrAla 48
DB 248 GGATATGGACGATAGAGATATTGATCTGAAA-----AGAGGCTTTCGT 292
QY 49 PheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTyr 68
DB 293 TTGTGGAAATTGAGATCCAGGATCCAGGATGAGATGATGCTGTGTATGATGCTGTGAGAAA 352
QY 69 AspPheAspGlyCysArgLeuArgValGluIleAlaHis-----Gly 82
DB 353 GAATC-----TGATGAAAGGTTACTATTGACATGCTAGGCTCGGTACGAGGT 406
QY 83 GlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSerArg--- 101
DB 407 GGAAGA-----CGTAGAGGACGATACTCTGACCGGTTTGTAGTAGTCGAGACCT 454
QY 102 -----AlaProSerArgArgSerAspTyrArgValLeuValThrGly 115
DB 455 CGAAATGATAGAGAAATCTCCACCTGTAAAGACAGAAATCGCTTATAGTTGAGAAAT 514
QY 116 LeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspVal 135
DB 515 TTATCTCAAGAGTCAGCTGGCAGGATCTCAAAGATTTCATGACAAAGCTGGGGAATA 574
QY 136 CysPheSerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsn 155
DB 575 ACGTTTGGGATGCACCGACCTAAATTAAT---GAAGGGGTGTTGAGTTCCCTCT 631
QY 156 TyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPhe 175
DB 632 TATGGTGACTTAAAGAAATGCTATTGAAAACTTCTGGAAGAGAAATA---AATGGGAGA 688
QY 176 SerSerAlaTyrIleArgValGluTyrGluSerArgSerValSerArgSerProAsp 195
DB 689 AAAATAAAATTAATTGAAGGACGACAAAGCCAGTAGGTCAAGAGCAGGCTCGATCC 748
QY 196 AspSerLysSerTyrArgSerArgSerArgGlyProSerCysSerTyrSerSer 215
DB 749 CGGACCCAGAGT---TCCTCTAGGTCTCGTAGCCGATCCCGTCCCGTAGTGCAGAACT 805

QY 216 LysSerArgSerValSerProAlaArgSerIleSerProArgSerArgProLeuSerArg 235
DB 806 TACAGCCGGTCA-----AGAGCAGGAGCAGGACCG 838
QY 236 SerArgSerLeuTyrSerSerValSerArgSer----- 246
DB 839 AGCCGGAGCAAGTCCCGTTCTGTAGGTCTCCCGTCTGCTGAGAGAGCCAGAAACGT 898
QY 247 GlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSerLysSerArgSerArg 266
DB 899 GGTCTTCAAGTAGATCTAAGTCTCCAGCATCTGTGATCGCCAGAGGTCCCGGTCCCGA 958
QY 267 SerArgSerAsnSer 271
DB 959 TCAAGGTCAGATCA 973

RESULT 3
US-09-313-294A-5971
; Sequence 5971, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalugudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5971
; LENGTH: 284
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350992H1
; NAME/KEY: unsure
; LOCATION: 70, 83, 238, 258-259, 283
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-5971

Alignment Scores:
Pred. No.: 5.12e-15 Length: 284
Score: 254.50 Matches: 60
Percent Similarity: 71.58% Conservative: 8
Best Local Similarity: 63.16% Mismatches: 21
Query Match: 17.54% Indels: 6
DB: 4 Gaps: 1

US-10-014-927-19 (1-279) x US-09-313-294A-5971 (1-284)
QY 39 LysIleProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAla 58
DB 3 AAGTCCCCCAAGACACCTGGTTATCTTTGTGAGTTTGAAGATCTCTGTGATCT 62
QY 59 AspAspAlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGlu 78
DB 63 GAGGAGGNAATTCGTCAACGNGATGGATACAACTTTGATGACACCGCTCTAAGAGTGGAG 122
QY 79 IleAlaHisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSer 98
DB 123 GCTGCTCATGGTGTAGAGTAATGCTTCTCCATGATGCTTCAAGTGGCTGGTGGCG 182
QY 99 Ala-----SerArgAlaProSerArgArgSer-AspTyrArgValLeuVal-ThrG 115
DB 183 GTGGTGGAGCACGTACGTGGTGTGTCGAGACACTCCAGAGTATCGTGTCTTGTCTACTG 242
QY 115 LysLeuProProSer-AlaSerTrpGlnAspLeuLysAsp 127
DB 243 GAATGCCTTCTTCTGNNATCAGGCGCAGGATTTAAGGAT 281

```

RESULT 4
US-09-976-594-960
; Sequence 960, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN CSA LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 960
; LENGTH: 2886
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. 6673549 988704.25
; NAME/KEY: unsure
; LOCATION: 17, 2848
; OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-960

Alignment Scores:
Pred. No.: 6.04e-10 Length: 2886
Score: 214.50 Matches: 68
Percent Similarity: 53.16% Conservative: 16
Best Local Similarity: 43.04% Mismatches: 53
Query Match: 14.78% Indels: 21
DB: 4 Gaps: 5

US-10-014-927-19 (1-279) x US-09-976-594-960 (1-2886)
Qy 123 GlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPhePro 142
Db 2108 AAGGATCTCAAGATTTCATGAGCAAGCTGGGAGATTAAGTTTCGGATGCACACCGA 2167
Qy 143 AspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAla 162
Db 2168 CCTAAATTAAAT---GAAGGGTGGTTGAGTTTGCCTCTTATGGTGACTTAAAGAAATGCT 2224
Qy 163 IleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal 182
Db 2225 ATTGAAAAAATCTTCGGAAGGAAATA---AATGGGAGAAAAATAAAATTAAATTGAAGGC 2281
Qy 193 ArgGluTyrGluSerArgSerValSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSer 202
Db 2282 AGCAAAAGGCACAGTAGGTCAAGAGCAGGTCTCGATCCCGGACCAAGAT---TCCTCT 2338
Qy 203 ArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSerPro 222
Db 2339 AGGTCCTAGCCGATCCCGTTCCTAGTAGTCGCAAACTTACACCCGGTCA----- 2389
Qy 223 AlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSer 242
Db 2390 -----AGAGCAGGAGCAGGAGCCGAGCCGAGCCGAGCCGAGTCCCGTTCT 2431
Qy 243 ValSerArgSer-----GlySerLeuLeuArgAlaGly 253
Db 2432 GTTAGTAGTCTCCCGTCCCTGAGAGAGCCAGAACCGTGGTTCTTCAAGTAGATCTAAG 2491
Qy 254 AspTyrIleSerGlnSerArgSerLysSerArgSerArgSerArgSerArgSerAsnSer 271
Db 2492 TCTCCAGCATCTGTGATCGCCAGAGGTCCCGGTCCCGATCAAGGTCCAGATCA 2545

RESULT 5
US-09-620-312D-524
; Sequence 524, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radjoje T.
; TITLE OF INVENTION: NO. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784C1P2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 524
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(866)
US-09-620-312D-524

Alignment Scores:
Pred. No.: 4.93e-10 Length: 955
Score: 208.00 Matches: 84
Percent Similarity: 43.65% Conservative: 44
Best Local Similarity: 28.67% Mismatches: 91
Query Match: 14.33% Indels: 74
DB: 4 Gaps: 12

US-10-014-927-19 (1-279) x US-09-620-312D-524 (1-955)
Qy 6 AsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAsp 25
Db 105 AACACCTCCTCTTCATCAGGAACCTCGCGGACGCCACCGCCTGAGGACTTGGCCGT 164
Qy 26 LeuPheTyrLysTyrGlyProIleValAlaPheLeuAspLeuLysIle-----ProPro 42
Db 165 GAGTTTGGTCGATATGGCCCTATAGACGTTTACATTCACCTTGACTTCTACACTGC 224
Qy 43 ArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIle 62
Db 225 CGCCCAAGAGGATTTCCTTATTTCAATTGAGATGTTTCGAGATGCTGAGATGCTCTT 284
Qy 63 TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLysLeuArgValGluIleAlaHisGly 82
Db 285 TATACTCTCAATAGAAGTGGGTATGTGCGCGCTCAGATGTAATACAGTTTGCACAGGT 344
Qy 83 GlyArgArg-----PheSerProSerVal 90
Db 345 GATCGCAAAACACCAAGCCCAATGAATCAAAAGAACGTCATCTCTTGTCTTCCAGTGT 404
Qy 91 AspArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgSerAspTyrArg 110
Db 405 CACAGGAGATCAAGAGCCCGCCAGCAAGAACCTCGAAGTAGA----- 449
Qy 111 ValLeuValThrGlyLeuProProSerAlaSerTrp-----GlnAspLeuLysAsp 127
```

Db 450 -----AGTTCTTCATGGGGAAGAAATAGAGGCGGTACAGC 485
Qy 128 HisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysGlyMet 147
Db 486 AGCTTAAGAGTCTCGACACAGCGCATTTCT 518
Qy 148 SerGlyValValAspTyrSerAsnTyrAspMetLysTyrAlaLysLysLeuAsp 167
Db 519 -----TATAGCCAGTCTAAATCTCGTTCCAAATCATTACCA 554
Qy 168 AlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal-----ArgGluTyr 185
Db 555 -----AGGCGGTCTACCTCAGCAAGCAGTCAAGACTCAAGAGAGGTTT 602
Qy 186 GluSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArg 205
Db 603 GGTCTTAGAGGCGGTACAGG-----TCCAGTCTTACAAAGAGGTTCCAG 650
Qy 206 SerArgGlyProSerCysSerTyrSerLysSerArgSerValSerProLysSerArgSer 225
Db 651 TCAATAGGA-----AAATCAGTCAAGTTTCACTCAAAAGCAG 689
Qy 226 IleSerProArgSerArgProLeuSerArgSerArgSerLeuTyrSer---SerValSer 244
Db 690 ACTAGCTCAGGAACAAA-----TCAAGTACATGGAAGACATTTCTCACTCAATAGCA 743
Qy 245 ArgSerGlySerLeuLeuArgAlaGlyAspTyrIleSerGlnSerArg-----260
Db 744 AGATCCCGGTGTAATCTCCCAAGGGTATACCAATCTCAAACTAAAGTACAAACAGCA 803
Qy 261 -----SerLysSerArgSerArgSerArgSerAsnSer 271
Db 804 AAGCANTCTCATTTTGGTGCACATTCAGATCTCGAAGT 842

RESULT 6
US-09-702-705-1719
; Sequence 1719, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1933
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1719
; LENGTH: 193
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-1719

Alignment Scores:
Pred. No.: 3,78e-09 Length: 193
Score: 187.50 Matches: 40
Percent Similarity: 69.57% Conservatives: 8
Best Local Similarity: 57.97% Mismatches: 12
Query Match: 12.92% Indels: 9
DB: 4 Gaps: 2

US-10-014-927-19 (1-279) x US-09-702-705-1719 (1-193)
Qy 53 GluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAspPheAspGly 72

Db 3 GAGACCCCCAGATGCAGAGGATGCTATTATTATGGAAGAAATGGTTATGATTATGGCCAG 62
Qy 73 CysArgLeuArgValGluIle-----AlaHisGlyGlyArgArgPheSerProSerVal 90
Db 63 TGTCGGCTTCGTGTGGAGTTCCCCAGGACTTATGAGAGTCTGGGTGGTGGCCCGGTGGT 122
Qy 91 AspArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgArgSerAspTyrArg 110
Db 123 GGGAGG-----AATGGGCTCTCTACAGAGATCTGATTTCCTCGA 161
Qy 111 ValLeuValThrGlyLeuProSer 119
Db 162 GTTCTTGTTCAGGACTTCCTCCGTCA 188

RESULT 7
US-09-736-457-1719
; Sequence 1719, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1719
; LENGTH: 193
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-736-457-1719

Alignment Scores:
Pred. No.: 3,78e-09 Length: 193
Score: 187.50 Matches: 40
Percent Similarity: 69.57% Conservatives: 8
Best Local Similarity: 57.97% Mismatches: 12
Query Match: 12.92% Indels: 9
DB: 4 Gaps: 2

US-10-014-927-19 (1-279) x US-09-736-457-1719 (1-193)
Qy 53 GluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAspPheAspGly 72

Db 3 GAGACCCCCAGATGCAGAGGATGCTATTATTATGGAAGAAATGGTTATGATTATGGCCAG 62
Qy 73 CysArgLeuArgValGluIle-----AlaHisGlyGlyArgArgPheSerProSerVal 90
Db 63 TGTCGGCTTCGTGTGGAGTTCCCCAGGACTTATGAGAGTCTGGGTGGTGGCCCGGTGGT 122
Qy 91 AspArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgArgSerAspTyrArg 110
Db 123 GGGAGG-----AATGGGCTCTCTACAGAGATCTGATTTCCTCGA 161
Qy 111 ValLeuValThrGlyLeuProSer 119
Db 162 GTTCTTGTTCAGGACTTCCTCCGTCA 188

RESULT 8
US-09-671-325-1719
; Sequence 1719, Application US/09671325
; Patent No. 6667154

Alignment Scores:
Pred. No.: 3,78e-09 Length: 193
Score: 187.50 Matches: 40
Percent Similarity: 69.57% Conservatives: 8
Best Local Similarity: 57.97% Mismatches: 12
Query Match: 12.92% Indels: 9
DB: 4 Gaps: 2

US-10-014-927-19 (1-279) x US-09-736-457-1719 (1-193)
Qy 53 GluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAspPheAspGly 72

Db 3 GAGACCCCCAGATGCAGAGGATGCTATTATTATGGAAGAAATGGTTATGATTATGGCCAG 62
Qy 73 CysArgLeuArgValGluIle-----AlaHisGlyGlyArgArgPheSerProSerVal 90
Db 63 TGTCGGCTTCGTGTGGAGTTCCCCAGGACTTATGAGAGTCTGGGTGGTGGCCCGGTGGT 122
Qy 91 AspArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgArgSerAspTyrArg 110
Db 123 GGGAGG-----AATGGGCTCTCTACAGAGATCTGATTTCCTCGA 161
Qy 111 ValLeuValThrGlyLeuProSer 119
Db 162 GTTCTTGTTCAGGACTTCCTCCGTCA 188

RESULT 8
US-09-671-325-1719
; Sequence 1719, Application US/09671325
; Patent No. 6667154

Alignment Scores:
Pred. No.: 3,78e-09 Length: 193
Score: 187.50 Matches: 40
Percent Similarity: 69.57% Conservatives: 8
Best Local Similarity: 57.97% Mismatches: 12
Query Match: 12.92% Indels: 9
DB: 4 Gaps: 2

US-10-014-927-19 (1-279) x US-09-736-457-1719 (1-193)
Qy 53 GluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAspPheAspGly 72

```
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1719
; LENGTH: 193
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-1719

Alignment Scores:
Pred. No.: 3,78e-09 Length: 193
Score: 187.50 Matches: 40
Percent Similarity: 69.57% Conservative: 8
Best Local Similarity: 57.97% Mismatches: 12
Query Match: 12.92% Indels: 9
DB: 4 Gaps: 2

US-10-014-927-19 (1-279) x US-09-671-325-1719 (1-193)

QY 53 GluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyrAspPheAspGly 72
Db 3 GAGACCCCGGAGATCAGAGGATGCTATTATGGAAGAAATGGTTATGATTCGCCAG 62
QY 73 CysArgLeuArgValGluile-----AlaHisGlyGlyArgArgPheSerProSerVal 90
Db 63 TGTGGCTTCGTGTGAGTCCCGAGGACTTATGAGGTCGGGTGGTGGTGGTGGTGGT 122
QY 91 AspArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgArgSerAspTyrArg 110
Db 123 GGAGG-----AATGGCCCTCCTACAAGAAGATCTGATTCGGA 161
QY 111 ValLeuValThrGlyLeuProSer 119
Db 162 GTTCTTGTTCAGGACTTCTCCGTC 188

RESULT 9
US-09-671-325-1719
; Sequence 867, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 867:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1015 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1049089
US-09-023-655-867

Alignment Scores:
Pred. No.: 2,36e-07 Length: 1015
Score: 179.00 Matches: 50
Percent Similarity: 53.98% Conservative: 11
Best Local Similarity: 44.25% Mismatches: 32
Query Match: 12.34% Indels: 20
DB: 4 Gaps: 6

US-10-014-927-19 (1-279) x US-09-023-655-867 (1-1015)

QY 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
Db 234 GTATTTCATCGGAGACATAATCCAGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGG 293
QY 29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla 48
Db 294 GGATATGACCGGATAGAGATATTGATCTGAAA-----AGAGGGCTTTGGT 338
QY 49 PheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAspGlyTyr 68
Db 339 TTGTGGAAATTGAGGATCCAGGATCCAGGATGATCTGTGTATGATGATGATGATGATG 398
QY 69 AspPheAspGlyCysArgLeuArgValGluLeuAlaHis-----Gly 82
Db 399 GAACCTC-----TGTAGTGAAGGGTTACTATTGACATGCTAGGGCTCGGTACGAGGT 452
QY 83 GlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSerArgAla 102
Db 453 GGAAGA-----GGTAGAGGACGATCTCTGACCGTTTGTAGT-----AGTCCGAGA 497

RESULT 10
US-09-621-976-3128
; Sequence 3128, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
```



```

; SEQ ID NO 3126
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 165..407
US-09-621-976-3128

Alignment Scores:
Pred. No.: 1-54e-06 Length: 479
Score: 165.00 Matches: 49
Percent Similarity: 54.13% Conservative: 10
Best Local Similarity: 44.95% Mismatches: 29
Query Match: 11.37% Indels: 21
DB: Gaps: 6

US-10-014-927-19 (1-275) x US-09-621-976-3128 (1-479)
QY 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 283
Db 180 GTATTTCATCGGAGAGACTAAATCCAGCGCCAGGAGAGACGCTGGAAAGATTCTTCAAG 239
QY 29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProProGlyTyrAla 48
Db 240 GGATATGGACGATAGAGATATTGATCTGAAA-----AGAGGCTTTGGT 284
QY 49 PheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyA-GAspGlyTyr 68
Db 285 TTGTGGAAATTTGAGATCCAGGATCCAGGATCCAGATGATGCTGTGTATGAGCTTGATGAAAA 344
QY 69 AspPheAspGlyCysArgLeuArgValGluLeuAlaHis-----G1 82
Db 345 GAATCT-----TGTAGTGAAAGGCTTACTATTGACATGCTAGGCGTCGTCACGAGG 398
QY 82 yGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSerArgal 102
Db 399 TGGAGA-----GGTAGAGGACGATACTCTTGACGCTTTAAST---AGTCGCGAG 443
QY 102 aProSerArgSerAspTyrArg 110
Db 444 ACCT-----CGAATGATAGACGG 462

RESULT 11
US-08-935-450-7
; Sequence 7, Application US/08935450
; Patent No. 5977311
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Meijia
; APPLICANT: Schulz, Vincent
; TITLE OF INVENTION: 53BP2 COMPLEXES
; FILE REFERENCE: 7934-054
; CURRENT APPLICATION NUMBER: US/08/935,450
; CURRENT FILING DATE: 1997-09-23
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1894
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12)..(1322)
US-08-935-450-7

Alignment Scores:
Pred. No.: 3-08e-05 Length: 1894
Score: 160.00 Matches: 96
Percent Similarity: 34.13% Conservative: 33
Best Local Similarity: 25.40% Mismatches: 127
Query Match: 11.03% Indels: 123
DB: Gaps: 15

```

Db 1041 AGACAGAAAGAGGCTTCCCTTCTATGAAAGG---GGGTACCTCTCCACGTGAT- 1096

Qy 254 AspTrpLeuSerGlnSerArgSerLysSerArgSerArgSerArgSerArgSer 271

Db 1097 TCCTACAGCAGTTCAAGCGCGGACGACCAAGAGGTGGTGGCGGTGGAGGAGC 1150

RESULT 12

US-09-338-123-7

; Sequence 7, Application US/09338123

; Patent No. 6627405

; GENERAL INFORMATION:

; APPLICANT: Nandabalan, Krishnan

; APPLICANT: Yang, Meijia

; APPLICANT: Schulz, Vincent

; TITLE OF INVENTION: 53BP2 COMPLEXES

; FILE REFERENCE: 7934-054

; CURRENT APPLICATION NUMBER: US/09/338.123

; CURRENT FILING DATE: 1999-06-22

; EARLIER APPLICATION NUMBER: 08/935,450

; EARLIER FILING DATE: 1997-09-23

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7

; LENGTH: 1894

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (12)..(1322)

US-09-338-123-7

Alignment Scores:

Pred. No.: 3.08e-05 Length: 1894

Score: 160.00 Matches: 96

Percent Similarity: 34.13% Conservative: 33

Best Local Similarity: 25.40% Mismatches: 127

Query Match: 11.03% Indels: 123

DB: 4 Gaps: 15

US-10-014-927-19 (1-279) x US-09-338-123-7 (1-1894)

Qy 9 lletyValGlyAsnLeuProGlyAspLeuArgLysCysGluValGluAspLeuPheTyr 28

Db 39 CTCCTTCATTGGTGGCTTAATACGGAAACAAATGAGAAAGCTCTTGAAGCAGTATTGGC 98

Qy 29 llyTyrGlyProIleValAspLeu-----LysIleProProArgProPro 45

Db 99 MAATATGGCAATAGTGGAAAGTACTCTTGATGAAAGACCGTGAACCAACAAATCAAGA 158

Qy 46 GlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArg 65

Db 159 GGATTGTCTTTGTCACTTTGAAGCCGACGACGCTAAGGATGACGCCAGAGACATG 218

Qy 66 AspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla----- 80

Db 219 AATGAAAGTCAATAGATGAAAGCCATCAAGGTGGAAACCAACCAACCATCATTT 278

Qy 81 HisGlyGlyArgPheSerProSerValAspArg----- 92

Db 279 GAAAGTGGTAGAGTGGACCGGCTCCACCTCAAGAGTAGAGGCGCTCCCAAGAGGTCTT 338

Qy 92 ----- 92

Db 339 AGAGGTGGAAGAGGAGGAAGTGGAGGAACCAAGGACCTCCCTCAAGGGAGGACACATG 398

Qy 93 -----TyrSerSerTyrSerAlaSer----- 100

Db 399 GATGACGGTGGATATTCATGAAATTTTACATGATCTTCCAGGGGACCACTCCAGTA 458

Qy 101 ---ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProPro--- 118

Db 459 AAAAGAGGACCAACCAAGAGAGTGGG-----GGTCTCTCTCTCTAAG 500

Qy 119 ---SerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPhe 137

Db 501 AGATCTGCACCTTCAGGACCAAGTTCGAGTAGCAGTGAATGGGAGGAGAGCTCTCTGTA 560

Qy 138 Ser-----GluValPheProAspArgLys 145

Db 561 TCACGTGGAAGAGATAGTTATGGAGGTCCACCTCGAAGGGAACCGCTCCCTCTCGTAGA 620

Qy 146 -----GlyMetSerGlyValValValAspTyrSerAsnTyr 156

Db 621 GATGTTTATTGTTCTCCAGAGAGATGATGGGTATTCTACTAAAGACAGACTATTTCAGAGCA 680

Qy 157 Asp-----AspMetLysTyrAlaIle 163

Db 681 GATTACCCCAAGTTCCTCGTGAATAGTAGAGATTATGCACCACCACGAGATTATATTAC 740

Qy 164 ArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSer----- 177

Db 741 COTGATTATGGTCATCTCCAGITTCACGTGATGACTATCCATCAAGAGAATATAGCGATAGA 800

Qy 178 ---AlaTyrIleArgValArgGluTyrGluSerArgSerValSerArgSerProAsp 196

Db 801 GATGGATATGGTCTGATCGTCTGACTATTCAAGTATCAAGTGGAGGTTCCTACAGAGAT 860

Qy 197 Ser---LysSerTyr-----ArgSerArgSerArgSerArgGlyProSerCysSer 212

Db 861 TCATATGAGAGTTATGGTAACCTCAGTAGTGTCCACCTACACAGAGGCCCCCGCACT 920

Qy 213 TyrSerSerLysSerArg----- 218

Db 921 TATGGTGAAGCAGTCGCTATGATTAACAGCAGCTCAGTCAGCGGATATGGTGAAGT 980

Qy 219 ---SerValSerProAlaArgSer-----IleSerProArgSerArgProLeu 233

Db 981 CGAGACAGTTACTCAAGCAGCGGAAAGTGTCTCTACTCAAGTGGTCGATCGGTGGC 1040

Qy 234 SerArgSerArgSerLeuTyrSerSerValSerArgSerGlySerLeuArgAlaGly 253

Db 1041 AGACAGAAAGAGGCGTCTCCCTTCTATGAAAGG---GGGTACCTCTCTCCACGTGAT- 1096

Qy 254 AspTrpIleSerGlnSerArgSerLysSerArgSerArgSerArgSerArgSer 271

Db 1097 TCCTACAGCAGTTCAAGCGCGGACGACCAAGAGGTGGTGGCGGTGGAGGAGC 1150

RESULT 13

US-09-613-182-14

; Sequence 14, Application US/09613182

; Patent No. 6294653

; GENERAL INFORMATION:

; APPLICANT: Mayfield, Stephen

; TITLE OF INVENTION: RNA BINDING PROTEIN AND BINDING SITE USEFUL FOR

; TITLE OF INVENTION: EXPRESSION OF RECOMBINANT MOLECULES

; FILE REFERENCE: SCR21778

; CURRENT APPLICATION NUMBER: US/09/613,182

; CURRENT FILING DATE: 2000-07-10

; PRIOR APPLICATION NUMBER: 09/341,550

; PRIOR FILING DATE: 1999-07-13

; PRIOR APPLICATION NUMBER: PCT/US98/00840

; PRIOR FILING DATE: 1998-01-16

; PRIOR APPLICATION NUMBER: 60/035,955

; PRIOR FILING DATE: 1997-01-17

; PRIOR APPLICATION NUMBER: 60/069,400

; PRIOR FILING DATE: 1997-12-12

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 14

; LENGTH: 1278

; TYPE: DNA

; ORGANISM: Chlamydomonas reinhardtii

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1272)

US-09-613-182-14

```

Alignment Scores:
Pred. No.: 5,6e-05 Length: 1278
Score: 154.50 Matches: 67
Percent Similarity: 41.36% Conservative: 55
Best Local Similarity: 22.71% Mismatches: 111
Query Match: 10.65% Indels: 62
DB: 3 Gaps: 10

US-10-014-927-19 (1-279) x US-09-613-182-14 (1-1278)

Qy 6 AsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAsp 25
Db 127 AACTCGTCGTACGTCGTCACCTGGAGAGGATGTACCGAGGCCACCGCTTCGAG 186
Qy 26 LeuPheTyrLysTyrGlyProIleValAspIleAspLeu-----LysIleProPro 42
Db 187 CTCTTCCTCCGTGTGGTCCCTGTGGCTCCATTCGTCGTGGCGGATCCGTCACGGCC 246
Qy 43 ArgProProGlyTyrAlaPheValGluPheGlu-----AspProArgAspAlaAsp 59
Db 247 CGTCGCTGGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCT 306
Qy 60 AspAlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIle 79
Db 307 CGCGCATGAGACCTGACTACCATGCTGTCGACGGCAAGCCTATGCGCATCATGTGG 366
Qy 80 AlaHisGlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAla 99
Db 367 TCGCAC-----372
Qy 100 SerArgAlaProSerArgSerAsp-----TyrArgValLeuValThrGlyLeuPro 117
Db 373 ---CGGACCTTCGGCCCAAGTGGCGCTGCGCAACATCTTCATCAAGAACCTGGAG 429
Qy 118 ProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPhe 137
Db 430 AAGACCATGACGCCAGGCCCTGACGACACCTTCGCGCTTCGCGCAAGATCTGTCC 489
Qy 138 SerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAsp 157
Db 490 TGAAGGTTCGCTACGACGCAAGCGGTGTGCAAGGGCTACGGCTTCGTGCACTTCGAG 549
Qy 158 AspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSer 177
Db 550 GACCAGCGCGTGCCTGCGCATTCGCGCATTCAGACC-----582
Qy 178 AlaTyrIleArgValArgGluTyrGlySerArgSerValSerArgSerProAspAspSer 197
Db 583 -----GTCAACACCAAGAGATGAGGCAAGATCGTACGTGGCCCTTCCTCAGAG 636
Qy 198 LysSerTyrArgSerArgSerArgSerArg-GlyProSerCysSerTyrSerSerLysSe 217
Db 637 CGCGCTACGCGCCCGGCAAGGACGCTTGTACACCAACGCTGTCG-----TC 684
Qy 217 rArgSerValSerProAlaArgSerIleSerProArgSerArg---ProLeuSerArgSe 236
Db 685 AAGAACTTCGCGCGGACATCGGCGACGACGAGCTGGGCAAGATGGCCACCGACCGCGC 744
Qy 236 rArgSer-----LeuTyrSer 241
Db 745 GAGATCACCGCGCGGTGTCATGAGGACGACCAAGGCGCGCAGCAAGGCTTCGCGCTTC 804
Qy 241 rSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSe 261
Db 805 ATCAACTTCAGGACCGCGAGT---CGCGCGCCCAAGTGGCTGG-----AGTACCTGAAC 855
Qy 261 rLysSerArgSerArgSerArgSerAsnSerProValSerPro 275
Db 856 GAGCGCGAGATGACGCGCAAGACCTGTACGCGCGCGCGCGCC 898

RESULT 14
US-09-613-182-5

```

```

; Sequence 5, Application US/09613182
; Patent No. 6294653
; GENERAL INFORMATION:
; APPLICANT: Mayfield, Stephen
; TITLE OF INVENTION: RNA BINDING PROTEIN AND BINDING SITE USEFUL FOR
; FILE REFERENCE: SCR2177S
; CURRENT APPLICATION NUMBER: US/09/613,182
; CURRENT FILING DATE: 2000-07-10
; PRIOR APPLICATION NUMBER: 09/341,550
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: PCT/US98/00840
; PRIOR FILING DATE: 1998-01-16
; PRIOR APPLICATION NUMBER: 60/035,955
; PRIOR FILING DATE: 1997-01-17
; PRIOR APPLICATION NUMBER: 60/069,400
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 2846
; TYPE: DNA
; ORGANISM: Chlamydomonas reinhardtii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (197)..(2065)
; US-09-613-182-5

Alignment Scores:
Pred. No.: 0.000174 Length: 2846
Score: 154.50 Matches: 67
Percent Similarity: 41.36% Conservative: 55
Best Local Similarity: 22.71% Mismatches: 111
Query Match: 10.65% Indels: 62
DB: 3 Gaps: 10

US-10-014-927-19 (1-279) x US-09-613-182-5 (1-2846)

Qy 6 AsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAsp 25
Db 260 AACTCGTCGTACGTCGTCACCTGGAGAGGATGTACCGAGGCCACCGCTTCGAG 319
Qy 26 LeuPheTyrLysTyrGlyProIleValAspIleAspLeu-----LysIleProPro 42
Db 320 CTCTTCCTCCGTGTGGCCCTGTGGCTCCATTCGCGTGGCGGATCCGTCACGGCG 379
Qy 43 ArgProProGlyTyrAlaPheValGluPheGlu-----AspProArgAspAlaAsp 59
Db 380 CGCTCGCTGGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACG 439
Qy 60 AspAlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIle 79
Db 440 CGCGCATGAGACCTGAACTACCATGCTGTAACGCGCAAGCCTATGCGCATCATGTGG 499
Qy 80 AlaHisGlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAla 99
Db 500 TCGCAC-----505
Qy 100 SerArgAlaProSerArgArgSerAsp-----TyrArgValLeuValThrGlyLeuPro 117
Db 506 ---CGGACCTTCGCGCCGCAAGTCGCGGCTCGGCAACATCTTCATCAAGAACCTGAC 562
Qy 118 ProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPhe 137
Db 563 AAGACCATGACGCCAAGGCCCTGACGACACCTTCGCGCTTCGCGCAAGATCTGTGCC 622
Qy 138 SerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAsp 157
Db 623 TGAAGGTTCGCTACGACGCAAGCGCGTGTGCAAGGGCTACGGCTTCGTGCACTTCGAG 682
Qy 158 AspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSer 177
Db 683 GACCAGCGCGTGGCGCATTCGCGCATTCAGACC-----715

```

```
OY 178 AlaTyrIleArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSer 197
Db 178 AlaTyrIleArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSer 197
OY 716 -----GTCACCAAGAGAGATTGAGGCGCAAGATCGGTACGTGGCCCTTCCAGAAG 769
Db 716 -----GTCACCAAGAGAGATTGAGGCGCAAGATCGGTACGTGGCCCTTCCAGAAG 769
OY 198 LysSerTyrArgSerArgSerArg-GlyProSerCysSerTyrSerSerLysSe 217
Db 198 LysSerTyrArgSerArgSerArg-GlyProSerCysSerTyrSerSerLysSe 217
OY 770 CGCGCTGACCGCCCGCCAGGAGCGTTGTACACCAACGCTGTCG-----TC 817
Db 770 CGCGCTGACCGCCCGCCAGGAGCGTTGTACACCAACGCTGTCG-----TC 817
OY 217 rArgSerValSerProAlaArgSerIleSerProArgSerArg---ProLeuSerArgSe 236
Db 217 rArgSerValSerProAlaArgSerIleSerProArgSerArg---ProLeuSerArgSe 236
OY 818 AAGAATCTGCCCGCCGACATCGCGCAGCAGCAGTGGCGCAGATGGCCACCGACACGGC 877
Db 818 AAGAATCTGCCCGCCGACATCGCGCAGCAGCAGTGGCGCAGATGGCCACCGACACGGC 877
OY 236 rArgSer-----LeuTyrSe 241
Db 236 rArgSer-----LeuTyrSe 241
OY 878 GAGATCACAGCGCGTGTGTCATGAGCAGCAGCAAGCGCGCAGCAAGGGCTTCGGCTTC 937
Db 878 GAGATCACAGCGCGTGTGTCATGAGCAGCAGCAAGCGCGCAGCAAGGGCTTCGGCTTC 937
OY 241 rSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSe 261
Db 241 rSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSe 261
OY 938 ATCAACTTCAAGGACCGCGAGT---CGGCGGCCAAGTGGCTGG-----AGTACCTGAAC 988
Db 938 ATCAACTTCAAGGACCGCGAGT---CGGCGGCCAAGTGGCTGG-----AGTACCTGAAC 988
OY 261 rLysSerArgSerArgSerArgSerArgSerProValSerPro 275
Db 261 rLysSerArgSerArgSerArgSerArgSerProValSerPro 275
OY 989 GAGCGCGAGATGAGCGGCAAGACCCCTGTACGCGCGCGCGGCC 1031
Db 989 GAGCGCGAGATGAGCGGCAAGACCCCTGTACGCGCGCGCGGCC 1031
```

RESULT 15

```
US-09-621-976-2159
; Sequence 2159, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S. J.Y.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2159
; LENGTH: 488
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 227..472
; NAME/KEY: misc feature
; LOCATION: 18,137
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-2159
```

```
Alignment Scores:
Pred. No.: 7,67e-05 Length: 488
Score: 146.50 Matches: 31
Percent Similarity: 57.53% Conservative: 11
Best Local Similarity: 42.47% Mismatches: 26
Query Match: 10.10% Indels: 5
DB: 4 Gaps: 1
```

US-10-014-927-19 (1-279) x US-09-621-976-2159 (1-488)

```
OY 7 ArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeu 26
Db 7 ArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeu 26
OY 230 AGGCGGTGTCTCGGSACTTCGACGACGACCCCGCAGTATGACCTCGAGCGCCTC 289
Db 230 AGGCGGTGTCTCGGSACTTCGACGACGACCCCGCAGTATGACCTCGAGCGCCTC 289
OY 27 PheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProArgProProGly 46
Db 27 PheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProArgProProGly 46
OY 290 TTCTCCAAGTACGCGGCCCATCGCGCGATCGACATGAAGCTA-----GGC 334
Db 290 TTCTCCAAGTACGCGGCCCATCGCGCGATCGACATGAAGCTA-----GGC 334
OY 47 TyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAsp 66
Db 47 TyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyrGlyArgAsp 66
OY 335 TATGCTTTTCACTTTGAAGATGACCGGTGACGACAGATGTATATAGCGGCTCTCGAA 394
Db 335 TATGCTTTTCACTTTGAAGATGACCGGTGACGACAGATGTATATAGCGGCTCTCGAA 394
```

```
OY 67 GlyTyrAspPheAspGlyCysArgLeuArgValGluIle 79
Db 395 AATGTGTCTTTTGGTTASAACCGGCGCAGGYTCTCTGTG 433
```

Search completed: March 26, 2004, 20:26:38
Job time : 117 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 26, 2004, 19:38:06 ; Search time 432 Seconds
(without alignments)
2404.444 Million cell updates/sec

Title: US-10-014-927-19
Perfect score: 1451
Sequence: 1 MSSRNRTYVGNLFGDIRK.....RSKSRSRNSPVPVWG 279

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10014927/runat_24032004_152927_10016/app_query.fasta_1.455
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -MOD=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-THR_MIN=0 -ALIGN=15 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-WAXLEN=200000000 -USER=US10014927@cgn_1.392@runat_24032004_152927_10016
-NCPUS=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TTIMEOUT=120 -WARM TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/2/pubpna/PTC_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/2/pubpna/PTC_US_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

RESULT 1					
US-09-938-842A-1199					
762	87.7	1272	9	US-09-938-842A-1199	Sequence 1199, Ap
762	87.7	1272	11	US-09-938-842A-1199	Sequence 1199, Ap
1209	61.1	886	12	US-10-424-599-134224	Sequence 134224, A
4044	60.3	875	13	US-10-014-927-18	Sequence 18, Appl
1263	58.9	854	12	US-10-425-114-25234	Sequence 125234, A
1383	58.8	853	12	US-10-424-599-112845	Sequence 112845, A
1187	58.2	845	12	US-10-425-114-20355	Sequence 20355, A
1320	58.2	845	12	US-10-425-114-24783	Sequence 24783, A
1690	56.8	824	12	US-10-425-114-25745	Sequence 25745, A
1438	54.4	790	12	US-10-424-599-83457	Sequence 83457, A
1491	53.0	768.5	12	US-10-424-599-83461	Sequence 83461, A
3299	40.1	542.5	12	US-10-006-285-513	Sequence 513, App
1714	37.4	517	14	US-10-171-581-9	Sequence 9, Appl
1069	35.6	517	9	US-09-880-107-3342	Sequence 3342, Ap
1162	35.6	517	9	US-09-925-300-661	Sequence 661, App
1475	35.6	517	15	US-10-264-049-273	Sequence 273, App
3697	30.4	440.5	13	US-10-002-600-58	Sequence 58, Appl
2167	29.7	431	15	US-10-439-703-38	Sequence 38, Appl
1534	28.8	418.5	13	US-10-044-090-268	Sequence 268, App
2212	28.7	416.5	9	US-09-925-297-142	Sequence 142, App
759	28.4	412.5	9	US-09-910-943-651	Sequence 651, App
1579	28.2	409	15	US-10-388-934-197	Sequence 197, App
446	28.1	407.5	14	US-10-060-036-52	Sequence 52, Appl
345	26.7	388	9	US-09-770-791-783	Sequence 783, App
746	23.7	344	9	US-09-910-943-530	Sequence 530, App
489	22.4	325	10	US-09-918-995-23144	Sequence 23144, A
715	20.4	295.5	14	US-09-910-943-212	Sequence 212, App
2601	20.4	289.5	14	US-10-101-510-334	Sequence 334, App
742	19.2	279	9	US-09-910-943-539	Sequence 539, App
282	19.0	275.5	11	US-09-864-408A-343	Sequence 343, App
307	18.8	273	9	US-09-294-093B-1314	Sequence 1314, Ap
1209	18.5	268.5	12	US-10-425-114-23865	Sequence 23865, A
1285	18.5	268.5	12	US-10-425-114-26648	Sequence 26648, A
1352	18.3	265	12	US-10-425-114-24065	Sequence 24065, A
1302	18.3	265	12	US-10-425-114-15470	Sequence 15470, A
1219	18.2	264	12	US-10-425-114-33401	Sequence 33401, A
1322	17.5	254	12	US-10-425-114-19768	Sequence 19768, A
1750	16.8	244	9	US-09-925-302-316	Sequence 316, App
1264	16.7	242	12	US-10-425-114-22637	Sequence 22637, A
1301	16.7	242	12	US-10-425-114-21964	Sequence 21964, A
2754	16.5	239	15	US-10-104-047-500	Sequence 500, App
3026	16.1	234	10	US-09-919-039-314	Sequence 314, App
432	15.9	231	9	US-09-920-300A-1284	Sequence 1284, Ap
432	15.9	231	13	US-10-033-528-1284	Sequence 1284, Ap
432	15.9	231	14	US-10-099-926-1284	Sequence 1284, Ap

ALIGNMENTS

TRANSGENIC PLANTS CONTAINING

ORGANISM: Arabidopsis thaliana
US-09-938-842A-1199

Alignment Scores:
Pred. No.: 9,88e-123 Length: 762
Score: 1272.00 Matches: 253
Percent Similarity: 91.67% Conservative: 0
Best Local Similarity: 91.67% Mismatches: 1
Query Match: 87.66% Indels: 23
DB: 9 Gaps: 2

US-10-014-927-19 (1-279) x US-09-938-842A-1199 (1-762)

```

QY 1 MetSerSerArgTTPAsnArgThrIleTyValGlyValAsnLeuProGlyAspIleArgLys 20
DB 1 ATGAGTAGCGGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 60
QY 21 CysGluValGluAspLeuPheTyValGlyValAsnLeuProGlyAspIleArgLys 40
DB 61 TGTGAGGTTGAAGATCTCTTCAAGTATGACCAATTTGGACATGATTTGAAGATT 120
QY 41 ProProArgProGlyTyValAlaPheValGluAspProArgAspAlaAsp 60
DB 121 CCACCGAGACCTCTCGTTATGCCCTTTCGAGTTTGAAGATCTCTCGATGCGAGCAT 180
QY 61 AlaIleTyValGlyValAspLeuPheTyValGlyValAsnLeuProGlyAspIleArgLys 80
DB 181 GCAATTTATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGCA 240
QY 81 HisGlyGlyArgArgPheSerProSerValAspArgTyValGlyValAsnLeuProGlyAsp 100
DB 241 CATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 300
QY 101 ArgAlaProSerArgArgSerArgTyValGluValLeuValThrGlyValProProSerAla 120
DB 301 CGTGACCTTCAAGACGCTCTGACACCGCGTGTGTGACCGGATTTACCGCTTCTGCT 360
QY 121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
DB 361 TCGTGACGAGACCTTAAAGCATCTGCTGATGAGTGGTGGTGGTGGTGGTGGTGGTGGT 420
QY 141 PheProAspArgLysGlyMetSerGlyValValAlaPheValGluAspLeuValGluVal 160
DB 421 TTCCCTGACGAGACCTTAAAGCATCTGCTGATGAGTGGTGGTGGTGGTGGTGGTGGT 480
QY 161 TyrAlaIleArgLysLeuAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 180
DB 481 TACGCAATPAGGAATCTGATGCCACTGATTTGAAATGCTTTCTGATGCTTATATA 540
QY 181 ArgValArgGluTyValGluSerArgSerValSerArgTyValGlyValValAlaPheSer 200
DB 541 CGGGTCAGGGAATATGATCGAGGAGTGTGAGTGAAGCCAGATGATTTCTAAAGCTAT 600
QY 201 ArgSerArgSerArgSerArgGlyProSerCysSerTyValSerLysSerArgSerVal 220
DB 601 AGAAGCAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 625
QY 221 SerProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeuTy 240
DB 627 TCACCTGCTAGATCCATTTCCCGCGCTTCACGCGCCCTTATGCTGTTCTCGCTTATAC 686
QY 241 SerSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArg 260
DB 687 AGCTCTGTCTCAAGG-----TCCCAATCAAGA 713
QY 261 SerLysSerArgSerArgSerArgSerAanSerProValSerProVal 276
DB 714 TCAAAATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAG 761

```

RESULT 2

US-09-938-842A-1199
; Sequence 1199, Application US/09938842A
; Publication No. US20040009476A9

GENERAL INFORMATION:

APPLICANT: Harper, Jeff
APPLICANT: Krepes, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 1199
LENGTH: 762
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-1199

Alignment Scores:

Pred. No.: 9,88e-123 Length: 762
Score: 1272.00 Matches: 253
Percent Similarity: 91.67% Conservative: 0
Best Local Similarity: 91.67% Mismatches: 1
Query Match: 87.66% Indels: 23
DB: 11 Gaps: 2

US-10-014-927-19 (1-279) x US-09-938-842A-1199 (1-762)

```

QY 1 MetSerSerArgTTPAsnArgThrIleTyValGlyValAsnLeuProGlyAspIleArgLys 20
DB 1 ATGAGTAGCGGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCG 60
QY 21 CysGluValGluAspLeuPheTyValGlyValAsnLeuProGlyAspIleArgLys 40
DB 61 TGTGAGGTTGAAGATCTCTTCAAGTATGACCAATTTGGACATGATTTGAAGATT 120
QY 41 ProProArgProGlyTyValAlaPheValGluAspProArgAspAlaAsp 60
DB 121 CCACCGAGACCTCTCGTTATGCCCTTTCGAGTTTGAAGATCTCTCGATGCGAGCAT 180
QY 61 AlaIleTyValGlyValAspLeuPheTyValGlyValAsnLeuProGlyAspIleArgLys 80
DB 181 GCAATTTATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGCA 240
QY 81 HisGlyGlyArgArgPheSerProSerValAspArgTyValGlyValValAlaPheSer 100
DB 241 CATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 300
QY 101 ArgAlaProSerArgArgSerArgTyValGluValLeuValThrGlyValProProSerAla 120
DB 301 CGTGACCTTCAAGACGCTCTGACACCGCGTGTGTGACCGGATTTACCGCTTCTGCT 360
QY 121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
DB 361 TCGTGACGAGACCTTAAAGCATCTGCTGATGAGTGGTGGTGGTGGTGGTGGTGGTGGT 420
QY 141 PheProAspArgLysGlyMetSerGlyValValAlaPheSerArgTyValGlyValValAla 160
DB 421 TTCCCTGACGAGACCTTAAAGCATCTGCTGATGAGTGGTGGTGGTGGTGGTGGTGGT 480
QY 161 TyrAlaIleArgLysLeuAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 180
DB 481 TACGCAATPAGGAATCTGATGCCACTGATTTGAAATGCTTTCTGATGCTTATATA 540
QY 181 ArgValArgGluTyValGluSerArgSerValSerArgTyValGlyValValAlaPheSer 200
DB 541 CGGGTCAGGGAATATGATCGAGGAGTGTGAGTGAAGCCAGATGATTTCTAAAGCTAT 600

```

Qy	1	MetSerSerArgTrpAsnA	gThrIleTyrValGlyAsnLeuProGlyAspIleArgLys	20
		::::	::::	
Db	73	ATGAGCGGGGGTCAAGCGCACAAT	TATGTGTGGCAATCTTCGTGCGACGTCGTGGTTG	132
Qy	21	CysGluValGluAspLeuPheTyr	LYenYrGlyProIleValAspIleAspLeuLysIle	40
Db	133	AGAAAGTAGAGGATCATCTTCTA	CAAGTATGGTCCTATTGTTGCATTTGATTGAACATA	192
Qy	41	ProProArgProProGlyTyrAla	pheValGluPheGluAspProArgAspAlaAspAsp	60
		::::	::::	
Db	193	CCTCCGAGACCACGAGGTATGCT	TTTGTAGAGTTTGAGGATGCTCGTGATGCTGAAGAT	252
Qy	61	AlaIleTyrGlyArgAspGlyTyr	AspPheAspGlyCysArgLeuArgValGluIleAla	80
		::::	::::	
Db	253	GCAATTCAATCAAGATGGTTACA	ATTTGGTTTCGTTACGAGTTGAACTTGCA	312
Qy	81	HisGlyGlyArgArgPheSerPro	SerValAspArgTyrSer-----SerSerTyrSer	98
		::::	::::	
Db	313	CATGGTGACCGGGATATTCATCA	TCACTACAGTACCGCTTACAGTACCTATAGCGGTGGCAGT	372

Qy 1 MetSerSerArgTrrAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgGlys 20
 908 ATGAGTAGCCGATGAATCGTACGATCTACGTTGGGAATTCGCTGGAGATATTCGCAAG 867
 Db

Db 72 ATGAGCAGCGCAACAGCGGTACCAATCTATGTAGGCAATCTCCCTGGGGACATCCGTGAG 131
Qy 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspLeuLysIle 40
Db 132 AGGAGGTTGAGGATCTCTTCAAGATATGCGGTATTTTGGATATGACTTGAATA 191
Qy 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60
Db 192 CTTCCGAGACCTCTCGATACGATCTGTTGATTTGAGATCCACGTGATCTGATGAT 251
Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuValGluIleAla 80
Db 252 GCAATTTATGGCGGTATGGGTATTAATCTGATGCTACAGGTGAGGTTGAATTAGCT 311
Qy 81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSer 100
Db 312 CATGGTGGCAGAGCGCAGCTTATTTCTTATGATCGTTCAAGCAGCTATGAGTGCATGC 371
Qy 101 Arg---AlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProSer 119
Db 372 CGTGGAGGTGTTTCTAGGCGCTCTGATTTCCGTTGTTATGTTCACTGCTACCTCATCG 431
Qy 120 AlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGlu 139
Db 432 GCATGTGGCAAGATCTGAAGGACCATGCGGCGCTGTTGATCTGTTCTCTGAT 491
Qy 140 ValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMet 159
Db 492 GTATCCGTCAGGCTGAGAACTATTTGGAATTTGATATACAAATTTATGACGATATG 551
Qy 160 LysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyr 179
Db 552 AATACCGGATTAGGAAGCTTGTGATCTACATGCTAGGATGCTATTTCAAGAGCATAT 611
Qy 180 IleArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSer 199
Db 612 ATCAGGCTGAGGAGATGATGCTAGATACGAGCAGAGAGC----- 653
Qy 200 TyrArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSer 219
Db 654 ---CGTAGCCACTCGTACTCTAGAACCCCGCAGCTACAGCAGGAGGAGTCCAAATCT 710
Qy 220 ValSerProAlaArgSerLysSerProArgSerArgProLeuSerArgSerArgSerLeu 239
Db 711 GTTTCTCAGTCACCTCA---TCTGTGATGAAGATCGTATCAAGATCTGATCCCA 767
Qy 240 TyrSerSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSer 259
Db 768 ATTTCTTCTCTTCTCATGCA-----AGATATGCGCAAGCCCT 806
Qy 260 ArgSerLysSerArgSerArgSerArgSer 269
Db 807 AGGAGCAGACGCGCATCCCGTTCTCGGTCT 836

RESULT 6
US-10-424-599-112845
; Sequence 112845, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kowalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21153223B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 112845
; LENGTH: 1383
; TYPE: DNA
; ORGANISM: Glycine max

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_72906C.1
US-10-424-599-112845
Alignment Scores:
Pred. No.: 6,82e-79 Length: 1383
Score: 853.00 Matches: 190
Percent Similarity: 69.38% Conservative: 23
Best Local Similarity: 61.89% Mismatches: 58
Query Match: 58.79% Indels: 36
Gaps: 6
DB: 12

US-10-014-927-19 (1-279) x US-10-424-599-112845 (1-1383)
Qy 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
Db 153 ATGAGTAGACGCTCGAGCAGAACTGTCTAGCTTGGGAATCTACCGGTGATATCCGTGAA 212
Qy 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspLeuLysIle 40
Db 213 AGAGAAGTTGAAGATTTGTTTCTGAAGTAGCGACACATAACACATCGACCTAAAGTT 272
Qy 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60
Db 273 CCCCCAAGACACCTGTTATGCAATTTGTAGAGTTTGAAGATGCTCAAGATGCTGAGGAT 332
Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuValGluIleAla 80
Db 333 GCATTCGTGTCGTGATGGCTATGATTTGATGCCACCGTTACGGGTGAGCTGCC 392
Qy 81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSer----- 98
Db 393 CATGTGGCGCTGGTCAATTCATCTCAAAAGATCGGCACATAGTTCACAGCAATGGCGGA 452
Qy 99 AlaSerArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProPro 118
Db 453 GGTGACGCTGGGTATCTAGGCGCTGTAATATCGTGTCTAGTCTACCTGGATTTGCCCTCT 512
Qy 119 SerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSer 138
Db 513 TCGGCATCTCGCAGGATCTTAAAGATCACAATGCGCAGGAGGAGGATTTGCTTTCT 572
Qy 139 GluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAsp 158
Db 573 CAAATTTTTCATGATGGAAGGGTACTACCGGTATTTGGATTTACACAAATTTATGATGAT 632
Qy 159 MetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAla 178
Db 633 ATGAAATATGCTATTAAAGAGCTTGATGACTCTGAGTTCCGGAATGCAATTTTCCAAGGGA 692
Qy 179 TyrIleArgValArgGluTyrGluSer---ArgSerValSerArgSerProAspAspSer 197
Db 693 TATGTTCGGGTGAGGGAATATGATTCAAAGCGGGATTTCTTTAGAGTCTTAGTATGCTGC 752
Qy 198 LysSerTyr-----ArgSerArgSerArgSerArgGlyProSerCysSerTyrSer 214
Db 753 CCATCTCATTTAGAGGAGGAGCTATAGCCGCTAGCGGTAGCCATAGTATACAGT 812
Qy 214 ----- 214
Db 813 CGGAGCCGAGGCAAAAGCAATCTCCAAAAGGAAAGTCTTCACAGCGTTACCTGCTAAA 872
Qy 215 -----SerLysSerArgSerValSerProAlaArgSerIleSer-----Pro 228
Db 873 TCTCAAAAGGAAAGTCTTCACAGCGTTCACCTGCTTAATTTCCATTCAGATCTCCCTCT 932
Qy 229 ArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerArgSerGlySer 248
Db 933 CGTTCTCGCTCAAGTCAAGTCCCGTCCGTTTATCAGATCAGGATCAGGTCACGGTCCAGGCT 992
Qy 249 LeuLeu-----ArgAlaGlyAspTrpIleSerGlnSerArgSerLysSerArg 264
Db 993 CCATTCCTCCCGGTAAATAAAGCCCCCAAAACCGCAGTGCTAGCAGGAGTCCCAAGTAGG 1052

```
QY 265 SerArgSerArgSerAsnSer 271
|||||
DB 1053 AGCAGGAGCAGGAGCAAGCT 1073

RESULT 7
US-10-425-114-20355
; Sequence 20355, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 20355
; LENGTH: 1187
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3180-041-E2_FLI
US-10-425-114-20355

Alignment Scores:
Pred. No.: 3,79e-78 Length: 1187
Score: 845.00 Matches: 186
Percent Similarity: 73.05% Conservative: 20
Best Local Similarity: 65.96% Mismatches: 44
Query Match: 58.24% Indels: 32
DB: 12 Gaps: 8

US-10-014-927-19 (1-279) x US-10-425-114-20355 (1-1187)
QY 1 MetSerSerArgTTPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
|||||
DB 73 ATGACGAGCGCTGGAGCGGACGATCTACGTCGGGAACCTCCCGCGGACATCCGGAG 132
|||||
QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLys 40
|||||
DB 133 AGGAGGTGGAGGATCTCTTACAGTATGCTAAATTTGTTGACATTGACCTGAAGTC 192
|||||
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60
|||||
DB 193 CCCCCAAGACCACCTGGTATGCTTTTGTGAGTTGAAGATCTCTCGTATGCTGAGGAG 252
|||||
QY 61 AlaIleTyrGlyValGluAspGlyTyrAspPheAspGlyCysArgLeuValGluIleAla 80
|||||
DB 253 GCATTTGCTGGACGGATGGATACACTTGTATGACACCTCTAAGATGGAGGCTGCT 312
|||||
QY 81 HisGlyGlyArgArgPheSerProSerValAspArg-----TyrSerSerSerTyr 97
|||||
DB 313 CATGTGTGTAGAGTAATGCTCTCTCGCATGATCGTTCAAGTGGCTTTGGTGGCGGTGT 372
|||||
QY 98 SerAlaSerArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuPro 117
|||||
DB 373 GGAGCAGCTGCTGGTGTGTGCGAGACACTCAGAGTATCGTGTCTGTCTACCTGGAGCTGCT 432
|||||
QY 118 ProSerAlaSerTyrGluAspLeuLysAspHisMetArgLysAlaGlyAspValCysPhe 137
|||||
DB 433 TCTTCTGCATCATGCGAGGATTTAAAGGATCATATGCGGAGGCTGGTGTGTTGTTTC 492
|||||
QY 138 SerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAsp 157
|||||
DB 493 TCTGAAGTATCGGAGGCGGTGGCCCGTAGAATTGGACTACACAAATATGAT 552
|||||
QY 158 AspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSer 177
|||||
```

RESULT 8

```
US-10-425-114-24783
; Sequence 24783, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 24783
; LENGTH: 1320
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3637-241-A12_FLI
US-10-425-114-24783
```

```
Alignment Scores:
Pred. No.: 4,35e-78 Length: 1320
Score: 845.00 Matches: 186
Percent Similarity: 73.05% Conservative: 20
Best Local Similarity: 65.96% Mismatches: 44
Query Match: 58.24% Indels: 32
DB: 12 Gaps: 8
```

US-10-014-927-19 (1-279) x US-10-425-114-24783 (1-1320)

```
QY 1 MetSerSerArgTTPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
|||||
DB 184 ATGACGAGCGCTGGAGCGGACGATCTACGTCGGGAACCTCCCGCGGACATCCGGAG 243
|||||
QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLys 40
|||||
DB 244 AGGAGGTGGAGGATCTGTTTACAGTATGTTAAATTTGTTGACATTGACCTGAAGTC 303
|||||
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60
|||||
DB 304 CCCCCAAGACCACCTGGTATGCTTTTGTGAGTTGAAGATCTCTCGTATGCTGAGGAG 363
|||||
```

Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db 364 GCAATTCCTGGACGGATGATACAACTTTGATGGACACCGCTAAAGAGTGGAGGTGCT 423
Qy 81 HisGlyGlyArgArgPheSerProSerValAspArg-----TyrSerSerTyr 97
Db 424 CATGGTGGTAGAGTAATGCTCTCCATGATCGTTCAAGTGGCTTTGGTGGCGGTGT 483
Qy 98 SerAlaSerArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuPro 117
Db 484 GGAGCAGCTCGTGGTGTGTCAGACACTCAGAGTATCGTTCCTGCTGAGCTGCT 543
Qy 118 ProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPhe 137
Db 544 TCTTCTGCATCGGACGAGTAAAGCATCATATCGGAAGCTGCGTGTATGTTGTTTC 603
Qy 138 SerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAsp 157
Db 604 TCTGAAGTGTATCGCGAAGCGGTGGCCACCGTAGGAATTGTGGACTACACAAATTATGAT 663
Qy 158 AspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSer 177
Db 664 GATATGAAATATGCTATTAAGAAAGCTGGATGATGATCTGAATTCAGGAACGCCCTTTGGCGA 723
Qy 178 AlaTyrIleArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSer 197
Db 724 GCCTATATAGGTGAAGATAT-----AACGGC 753
Qy 198 LysSerTyrArgSerArgSerArgSerArgGlyProSerCysSerTyrSerLysSer 217
Db 754 AAACGTGGTGGCTCCTACTCAGCTAGCGGAAGCCCAAGTCGTAGTTAC---AGCAAAAGC 810
Qy 218 ArgSerValSerProAlaArgSerIleSerProArgSerArg-----ProLeuSer 234
Db 811 AGGAGTCCGAGTAA-----TCACCAGGACTCGCGTTCATCATCTAGATCC 858
Qy 235 ArgSerArgSerLeuTyrSerSerValSerArgSerGlySerLeuLeuAlaGlyAsp 254
Db 859 CGGTCAAGATCTGT---TCTTCTCGTTCAAGTCCCA-----894
Qy 255 TrpIleSerGlnSerArgSerLysSerArgSerArgSerArgSerAsnSerPro---Val 273
Db 895 -----TCAAAAGGACGCTTCTCATCAAGATCACCAGCAGATCGAAATCTCCTAATGTT 948
Qy 274 SerPro 275
Db 949 TCTCCA 954

RESULT 9
US-10-425-114-25745
; Sequence 25745, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25745
; LENGTH: 1690
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3959-003-F10_FLI
US-10-425-114-25745

Alignment Scores:
Pred. No.: 9,06e-76 Length: 1690
Score: 824.00 Matches: 187
Percent Similarity: 58.50% Conservative: 23
Best Local Similarity: 52.09% Mismatches: 45
Query Match: 56.79% Indels: 104
DB: 12 Gaps: 5

US-10-014-927-19 (1-279) x US-10-425-114-25745 (1-1690)

Qy 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyValAsnLeuProGlyAspIleArgLys 20
Db 145 ATGAGCAGCGGCAACACCGCTACCATCTATGTAGCAATCTCCCTGGGAGCATCGTGAG 204
Qy 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLysIle 40
Db 205 AGGGAGGTGGAGATCTCTTCTACAAGTATGCCGTATTTTGGATATTGACTTGAATAA 264
Qy 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60
Db 265 CCTCCGAGACCTCCTCGATACGCATTCGTTGAGTTTGAGGATCCACGCTGATGCTGAT 324
Qy 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
Db 325 GCAATTTATGCCGTGATGGGTATAACTTTGATGGCTACAGGTTGAGGTTGAATTAGCT 384
Qy 81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSer 100
Db 385 CATGTTGCACAGCGCCAGCTCTTATCTTATGATCGTTCAAGCAGCTATAGCAGTGCATGC 444
Qy 101 Arg---AlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProSer 119
Db 445 CGTGGAGGTGTTTCTAGCGCTCTGATTTCGTTGTTGTTGTTGTTGTTGTTGTTGTTG 504
Qy 120 AlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGlu 139
Db 505 GCATCGTGGCAAGATCTGAGGACACATCGCGCGCGCTGTTGTTGTTGTTGTTGTTGTTG 564
Qy 140 ValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMet 159
Db 565 GTATACCTGAGGCTGGAGAACTATTGGAATTTGGATTATACAATATTAGCAATATG 624
Qy 160 LysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyr 179
Db 625 AAATACGCGATTAGGAAGCTTGATGACTCAGCTCAGATTCAGAAATGCAATTTTCAAG 684
Qy 180 IleArgValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSer 199
Db 685 ATCAGGTGAGGAGTATGATGCTAGATCAGCAAGCAGAGC-----726
Qy 200 TyrArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSer 219
Db 727 ---CGTAGCCACTCGTACTCTAGAAGCCCGAGCTACAGCAGGACGAGGAGTCCAAATCT 783
Qy 220 ValSerProAlaArgSer-----225
Db 784 GTTCTCAGTCACCCCTCATCTCTGGAATGAAAGGTTTGTGTTTTCCTAGATTGGC 843
Qy 225 -----225
Db 844 CATCGCGCATCTCGGGGCTGCATGTATTAATACTGATGTTTGTGAGAAATCGGAGTGA 903
Qy 225 -----225
Db 904 TGGTTTGGGACTTTCGATAGTAGATGGAATTGTGATGATGATGATGATGATGATGATG 963
Qy 226 IleSerProArgSer-----230
Db 964 ATCAGTCTAGGACGACAGATAGGCTGACAGTAAGAAAATGTGAAGAAAGCTTGTGGTGT 1023
Qy 230 -----230

184 uTyrGluSerArgSerValSerArgSerProAspSerLysSerTyrArgSerArgSer 204
877 TGATGGCCCAAGATCCAAAGTTATGGAAGATTCGATCTCGAAGC---CGTAGTCTAG 933
204 rArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaAr 224
934 CAGAGCCGTAGCAGAGC-----AACAGCAGGAGTCGCAGTTACTCCCCAAGGAG 984
224 gSerile---SerProArgSerArgProLeu---SerArgSerArgSer 238
985 AAGCAGAGGATCACCAGCTATTCTCCCGCTCATAGCAGATCTCGCTCT 1033

RESULT 14
US-09-880-107-3342
; Sequence 3342, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-NO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIORITY FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3342
; LENGTH: 1069
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U30825
US-09-880-107-3342

Alignment Scores:
Pred. No.: 3,77e-44 Length: 1069
Score: 517.00 Matches: 117
Percent Similarity: 67.91% Conservative: 29
Best Local Similarity: 54.42% Mismatches: 42
Query Match: 35.63% Indels: 28
DB: 9 Gaps: 6

US-10-014-927-19 (1-279) x US-09-880-107-3342 (1-1069)

Qy 9 lIeTyrValGlyAenLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
Db 98 ATCTACGTGGGAACTTCACCGACGTCGCGGAGAGGACTTGGAGGACCTGTTCTAC 157
Qy 29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProGly-TyrAl 48
Db 158 AAGTACGGCCGATCCGCGAGATCGAGCTCAAGAA-CGGCAGCGCTCGTGGCCCTCGC 216
Qy 48 aPheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTy 68
Db 217 CTTCGTGGCTTCGAGGACCCCGCAGATCGAGATGCTATTTATGGAAGAATGGTTA 276
Qy 68 rAspPheAspGlyCysArgLeuValGluLe-----AlaHisGlyGlyArgArgPh 86
Db 277 TGATTATGCCAGTGTGCGCTTCGTGGAGTTCGCCAGGACTTATGAGGTGGGGTGG 336
Qy 86 eSerProSerValAspArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgAr 106
Db 337 GTGGGCCCGTGGGGAGG-----AATGGGCTCTCTACAGAAG 375
Qy 106 gSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLy 126
Db 376 ATCTGATTTCCGAGTCTTGTTTCAGGACTTCCTCCGTGAGGAGCTCGCAGGACCTGAA 435

184 uTyrGluSerArgSerValSerArgSerProAspSerLysSerTyrArgSerArgSer 204
877 TGATGGCCCAAGATCCAAAGTTATGGAAGATTCGATCTCGAAGC---CGTAGTCTAG 933
204 rArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaAr 224
934 CAGAGCCGTAGCAGAGC-----AACAGCAGGAGTCGCAGTTACTCCCCAAGGAG 984
224 gSerile---SerProArgSerArgProLeu---SerArgSerArgSer 238
985 AAGCAGAGGATCACCAGCTATTCTCCCGCTCATAGCAGATCTCGCTCT 1033

RESULT 14
US-09-880-107-3342
; Sequence 3342, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-NO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIORITY FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3342
; LENGTH: 1069
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U30825
US-09-880-107-3342

Alignment Scores:
Pred. No.: 3,77e-44 Length: 1069
Score: 517.00 Matches: 117
Percent Similarity: 67.91% Conservative: 29
Best Local Similarity: 54.42% Mismatches: 42
Query Match: 35.63% Indels: 28
DB: 9 Gaps: 6

US-10-014-927-19 (1-279) x US-09-880-107-3342 (1-1069)

Qy 9 lIeTyrValGlyAenLeuProGlyAspIleArgLysCysGluValGluAspLeuPheTyr 28
Db 98 ATCTACGTGGGAACTTCACCGACGTCGCGGAGAGGACTTGGAGGACCTGTTCTAC 157
Qy 29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProGly-TyrAla 48
Db 158 AAGTACGGCCGATCCGCGAGATCGAGCTCAAGAA-CGGCAGCGCTCGTGGCCCTCGC 216
Qy 48 aPheValGluPheGluAspProArgAspAlaAspAspAlaIleTyrGlyArgAspGlyTy 68
Db 217 CTTCGTGGCTTCGAGGACCCCGCAGATCGAGATGCTATTTATGGAAGAATGGTTA 276
Qy 68 rAspPheAspGlyCysArgLeuValGluLe-----AlaHisGlyGlyArgArgPh 86
Db 277 TGATTATGCCAGTGTGCGCTTCGTGGAGTTCGCCAGGACTTATGAGGTGGGGTGG 336
Qy 86 eSerProSerValAspArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgAr 106
Db 337 GTGGGCCCGTGGGGAGG-----AATGGGCTCTCTACAGAAG 375
Qy 106 gSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLy 126
Db 376 ATCTGATTTCCGAGTCTTGTTTCAGGACTTCCTCCGTGAGGAGCTCGCAGGACCTGAA 435

Qy 126 sAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysG1 146
Db 436 GSATCATCGAGAGATGGGATCTCTGTTATGCTGATGTCAGAGAGATGGAGTGGG 495
Qy 146 yMetSerGlyValValAspTyrSerAsnTyrAspMetLysTyrAlaIleArgLysLe 166
Db 496 GATG-----GTGAGTATCTCAGAAAGAGACATGGATATGCCCTCGTAACT 546
Qy 166 uAspAlaThrGluPheArg---AsnAlaPheSerSerAlaTyrIleArgValArgGluTy 185
Db 547 GGATGACACCAAAATTCGGCTCTCATGAGGCTGAACCTCTCTACATCCGAGTT----- 598
Qy 185 rGluSerArgSerValSerArgSerProAspAspSerLysSerTyr-----ArgSe 202
Db 599 -----TATCCTGAGAGAGACACCGACTATGGCTACTCAGGTC 636
Qy 202 rArgSerArgSerArgGlyProSerCysSerTyrSerSerLys 216
Db 637 TCGGTCTGGGTCAAGGGCGCGTGACTCTCCATCAACAAAGCAGG 679

RESULT 15
US-09-925-300-661
; Sequence 661, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 661
; LENGTH: 1162
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1155)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-661

Alignment Scores:
Pred. No.: 4.2e-44 Length: 1162
Score: 517.00 Matches: 117
Percent Similarity: 67.91% Conservative: 29
Best Local Similarity: 54.42% Mismatches: 42
Query Match: 35.63% Indels: 28
DB: 9 Gaps: 6

US-10-014-927-19 (1-279) x US-09-925-300-661 (1-1162)
Qy 9 IleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluValGluAspLeupheTyr 28
Db 143 ATCTACGTGGGAACCTTCGACCGAGTCGGCGAGAGAGACTTGGAGGACCTGTCTTAC 202
Qy 29 LysTyrGlyProIleValAspIleAspLeuLysIleProProArgProGly-TyrAl 48
Db 203 AAGTACGGCGCATCCGCGAGATCGAGCTCAAGAA-CCGGCAGCGCCTCGTGCCTTCGC 261
Qy 48 aPheValGluPheGluAspProArgAspAlaIleTyrGlyArgAspGlyTy 68
Db 262 CTTCTGCGCTTCAGACCCCGAGATCGAGAGATGCTATTATGGAAGAAATGTTA 321
Qy 68 rAspPheAspGlyCysArgLeuArgValGluIle-----AlaHisGlyGlyArgArgPh 85
Db 322 TGATTATGGCCAGTGTGGCTTCGTGTGGAGTTCCCGCAGGACTATGAGGTCGGGTGG 381

Qy 86 eSerProSerValAspArgTyrSerSerSerTyrSerAlaSerArgAlaProSerArgAr 106
Db 382 GTGCCCCCGTGGTGGAGG-----AATGGGCTCTCTACAAAGAG 420
Qy 106 gSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrpGlnAspLeuLy 126
Db 421 ATCTGATTTCGGAGTTCTTGTTCAGGACTTCCTCCGTCAGGCAGCTGGCAGGACCTGAA 480
Qy 126 sAspHisMetArgLysAlaGlyAspValCysPheSerGluValPheProAspArgLysG1 146
Db 481 GGATCATCATCGAGAGAGCTGGGATGCTCTGTTATGCTGTGTCAGAGAGATGGAGTGGG 540
Qy 146 yMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAlaIleArgLysLe 166
Db 541 GATG-----GTGAGTATCTCAGAAAGAGACATGGAATATGCCCTGCGTAACT 591
Qy 166 uAspAlaThrGluPheArg---AsnAlaPheSerSerAlaTyrIleArgValArgGluTy 185
Db 592 GGATGACACCAAAATTCGGCTCTCATGAGGCTGAACCTTCCTACATCCGAGTT----- 643
Qy 185 rGluSerArgSerValSerArgSerProAspAspSerLysSerTyr-----ArgSe 202
Db 644 -----TATCCTGAGAGAGACACCGACTATGGCTACTCAGGTC 681
Qy 202 rArgSerArgSerArgGlyProSerCysSerTyrSerSerLys 216
Db 682 TCGGTCTGGGTCAAGGGCGCGTGACTCTCCATCAACAAAGCAGG 724

Search completed: March 26, 2004, 21:34:11
Job time : 441 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 26, 2004, 18:12:15 ; Search time 2829 Seconds
(without alignments)
2945.051 Million cell updates/sec

Title: US-10-014-927-19
Perfect score: 1451
Sequence: 1 MSSRNRTIYVGNLPGDIRK.....RSKRSRRSRNSPVPVWISG 279

Scoring table:
BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Ygapext 0.5
Ygapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2/USPTO.spool/US10014927/runat_24032004_152926_9909/app.query.fasta_1.455
-DB=EST -Qfmt=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10014927@cgn_1_1_3437@runat_24032004_152926_9909 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estcom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_plg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	992.5	68.4	604	9	AV826310 AV826310
2	947	65.3	650	14	CB258427
3	881	60.7	929	14	CK272077
4	880.5	60.7	948	14	CK270575
5	851	58.6	935	14	CK244849
6	845	58.2	1255	11	AY108591
7	840.5	57.9	864	29	CG848083
8	839	57.8	877	14	CK267541
9	839	57.8	910	14	CK250556
10	839	57.8	921	14	CK256886
11	834.5	57.5	896	14	CK275124
12	829.5	57.2	874	29	CG848162
13	829.5	57.2	903	14	CK298330
14	829.5	57.2	930	14	CK285308
15	829	57.1	884	14	CK265450
16	827.5	57.0	779	13	BQ802976
17	822.5	56.7	833	14	CK273791
18	815.5	56.2	817	14	CK253962
19	814	56.1	839	14	CK252379
20	812	56.0	933	14	CK284417
21	811	55.9	933	14	CK269749
22	811	55.9	949	14	CK264971
23	809	55.8	897	14	CK298535
24	809	55.8	974	14	CK254408
25	808.5	55.7	1106	14	CK206059
26	807	55.6	989	14	CK270936
27	807	55.6	999	14	CK253886
28	805.5	55.5	806	12	BM109633
29	805	55.5	926	14	CK250645
30	804	55.4	809	14	CD879553
31	803	55.3	919	14	CK259499
32	802	55.3	852	14	CD439565
33	799	55.1	913	14	CK259500
34	799	55.1	924	14	CK264893
35	796.5	54.9	667	12	BG451868
36	796.5	54.9	812	14	CK272649
37	795	54.8	675	12	BG648253
38	795	54.8	840	14	CB650587
39	792	54.6	877	14	CK267276
40	792	54.6	900	14	CK262606
41	792	54.6	927	14	CK271717
42	789	54.4	627	12	BG135885
43	787	54.2	877	12	BG444501
44	785.5	54.1	736	14	CA802008
45	784	54.0	647	12	BI931752

ALIGNMENTS

RESULT 1
AV826310
LOCUS AV826310
DEFINITION AV826310 RAFL8 Arabidopsis thaliana CDNA clone RAFL08-12-105 5', linear EST 01-APR-2002
ACCESSION AV826310.1 GI:19868370
VERSION AV826310
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.

REFERENCE 1 (bases 1 to 604)
 AUTHORS Seki,M., Narusaka,M., Iehida,J., Kamiya,A., Satou,M., Nakajima,M.,
 Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
 Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
 and Shinozaki,K.
 TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
 JOURNAL Unpublished (2002)
 COMMENT Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: msekic@cc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
 and XhoI was ligated to modified lambda FIC-1 vector (Carninci et
 al., submitted for publication) digested with BamHI and SalI. This
 clone is in a modified pluscript vector. Please visit our web
 site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
 details.

FEATURES

source
 1..604
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAF108-12-105"
 /dev_stage="rosette plants"
 /lab_host="DH10B"
 /clone_lib="RAF108"
 /note="Site 1: BamHI; Site 2: SalI; subjected to
 dehydration-treated (1, 2, 5, 10, 24 hr)"
 Alignment Scores:
 Pred. No.: 4,51e-63 Length: 604
 Score: 992.50 Matches: 197
 Percent Similarity: 94.2% Conservative: 1
 Best Local Similarity: 93.81% Mismatches: 11
 Query Match: 68.40% Indels: 11
 DB: 9 Gaps: 1
 US-10-014-927-19 (1-279) x AV826310 (1-604)

QY 62 IleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHis 81
 DB 3 ATTTATGACGTGATGTTATGATTTGATGGTGTGACCTCGGTTGAGATTGCACAT 62
 QY 82 GlyGlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSerArg 101
 DB 63 GGTGGTCGTAGATTTCACCATCATCAGTTGATAGGTACAGCAGCAGCTACAGTCGAGCCGT 122
 QY 102 AlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSer 121
 DB 123 GCACCTTCAGACGCTCTGACTACCATCGCTGTGTGACCGGATTACCGCTTCGTTCG 182
 QY 122 TrpGlnAspLeuLeuAspHisMetArgLysAlaGlyAspValCysPheSerGluValPhe 141
 DB 183 TGGCAGGACCTTACGATCAGTGGCAAGCTGAGATGCTGCTCTCTGAAGTTTC 242
 QY 142 ProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyr 161
 DB 243 CCTGACCGTAAAGGATGCTGGGGTTGTGGATTATATACCACTATGATGATGATGATAC 302
 QY 162 AlaIleArgLysLeuAlaThrGluPheArgAenAlaPheSerSerAlaTyrIleArg 181
 DB 303 GCANATAGGAACCTTGATGCCACTGAATTCGAATGTTTCTCTAGTGCCTATATACGG 362
 QY 182 ValArgGluTyrGluSerArgSerValSerArgSerProAspAspSerLysSerTyrArg 201
 DB 363 GTGAGGGATATAGTCGAGGAGTGTGATGCGAGCCAGATGATCTTAAAGCTATAGA 422
 QY 202 SerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSer 221

DB 423 AGCAGGAGTCCGAGCCGTGTCACAGCTGTAGCTATAGTACGAGCAGGAGTGTGTCA 482
 QY 222 ProAlaArgSerLeSerProArgSerArgProLeuSerArgSerArgSerLeuTyrSer 241
 DB 483 CTTGCTAGATCCATTTCCTCCGCGTTCACGCCCTTAGTCTGCTCGCTATACAGC 542
 QY 242 SerValSerArgSerGlySerLeuLeuArgAlaGlyAspTyrPheSerGlnSerArgSer 261
 DB 543 TCTGTCTCAAGG-----TCCCATCAAGATCA 569
 QY 262 LysSerArgSerArgSerArgSerAsnSer 271
 DB 570 AAATCAAGACACAGATCAAGATCGAATTCT 599
 CB258427 650 bp mRNA linear EST 06-NOV-2003
 28-E012735-014-002-H16-T7R MPZ-ADIS-014 Arabidopsis thaliana cDNA
 clone MPZp77H162Q 5-PRIME, mRNA sequence.
 ACCESSION CB258427
 VERSION CB258427.1 GI:32883200
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 650)
 AUTHORS Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,
 Michell-Olds,T. and Weishaar,B.
 TITLE Large-scale identification and analysis of genome-wide
 single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
 JOURNAL Genome Res. 13 (6), 1250-1257 (2003)
 MEDLINE 22683290
 PUBMED 12799357
 COMMENT Contact: Weishaar B
 ADIS DNA core facility at MPZ
 Max-Planck-Institute for Plant Breeding Research
 Carl-von-Linne Weg 10, 50829 Koeln, Germany
 Fax: 00492215062851
 Email: weishaar@mpiz-koeln.mpg.de
 Insert Length: 650 Std Error: 0.00
 Plate: 2 row: H column: 16
 Seq primer: T7R; CTAATACGACTCTACTATAGGA.
 Location/Qualifiers
 1..650
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /cultivar="Cape Verde Island (Cvi)"
 /db_xref="GABI:591109"
 /db_xref="taxon:3702"
 /clone="MPZp77H162Q"
 /tissue_type="whole plant"
 /dev_stage="adult plant, mixed stresses"
 /lab_host="E. coli TOP10"
 /clone_lib="MPZ-ADIS-014"
 /note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; cDNA
 library from Arabidopsis thaliana, accession Cvi; six
 weeks old total plants grown under long-day conditions in
 soil, whole adult plants were treated for 24 hours with
 different stresses, (1) at 4 Grad C in the dark, (2), at
 37 Grad C in the dark, (3) lying in the lab after removing
 from soil, (4) in the greenhouse after wounding leaves
 with a forceps, (5) in the lab watering with a 150 mM
 NaCl solution, (6) at 26 Grad C in the light/UV; equal
 quantities of stressed plant material were pooled; library
 was made at the Max-Planck-Institute for Plant Breeding
 Research, Cologne, Germany; cloning sites SalI-NotI,
 primer sites and orientation:
 T7-SalI-CCAGCGTCCG-5prime-cDNA-polyA-CC-NotI-SP6; Note:
 Sequencing granted in the context of the GABI Arabidopsis
 Verbund I: Genetic Diversity, 'Establishment of

high-efficiency SNP-based mapping tools and development of methods for genome-wide mutation detection' PI: Bernd Weishaar Sequence submission managed by RZPD/GABI-Primary database: <http://gabi.rzpd.de>. This clone is available from RZPD; contact RZPD (clone@rzpd.de) for further information."

ORIGIN

Alignment Scores:
Pred. No.: 1,08e-59 Length: 650
Score: 947.00 Matches: 177
Percent Similarity: 98.33% Conservatives: 0
Best Local Similarity: 98.33% Mismatches: 3
Query Match: 65.27% Indels: 0
DB: 14 Gaps: 0

US-10-014-927-19 (1-279) x CB258427 (1-650)

QY 1 MetSerSerArgTrrpAenArgThrileTyrValGlyAsnLeuProGlyAspIleArgLys 20
DB 111 ATGAGTGGCGATGGGAATCGTACGATCTACGTTGGGAATTCCTGGGAGATATCGCAAG 170
QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40
DB 171 TGTGAGTTGAGATCTCTTACAGATATGACCAATTTGTGGACATTTGATTTGAGATT 230
QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
DB 231 CCACCGAGACCTCTGTTATGCTTTGTGCGATTTGAAGATCTCTGATGCACACGAT 290
QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
DB 291 GCAATTTATGGAGTGTATGATGATTTTGTGATGTTGCGTCTGCGGTTGAGATTGCA 350
QY 81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSer 100
DB 351 CATGGTGGTGTAGATTTTACCATCAGGTGTAGTACAGCAGGCTACAGTCCGAGC 410
QY 101 ArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120
DB 411 CGTGACCTCTTAGACGCTCTGACTATCGGTCTGTGTACCGGATTTACCGCTTCTGCT 470
QY 121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140
DB 471 TCGTGCAGACCTTAAGGATCACATGCGAAAGCTGGAGATGTCTGCTTCTCTGAATT 530
QY 141 PheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLys 160
DB 531 TTCCCTGACCGTAAGGACATGTCTGGGGTGGGGATATAGCAACTATGATGATATGAAG 590
QY 161 TyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIle 180
DB 591 TAGCCAATAAGGAACCTTGATGCCACTGAATTTGGAATGCTTTCTCTAGTCTTATATA 650

RESULT 3

CK272077
LOCUS 929 bp mRNA linear EST 12-DEC-2003
DEFINITION EST718155 potato abiotic stress cDNA library Solanum tuberosum cDNA clone POAD303 5' end, mRNA sequence.

ACCESSION

CK272077

VERSION

CK272077.1 GI:39829055

KEYWORDS

EST.

SOURCE

Solanum tuberosum (potato)

ORGANISM

Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.

REFERENCE

1 (bases 1 to 929)

AUTHORS

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

TITLE

Generation of ESTs from abiotic stressed potato tissue

JOURNAL

Unpublished (2003)

COMMENT

Other ESTs: EST718156

Contact: Robin Buell

The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GfG ACA CTA TAG.

FEATURES

source

1. 929
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POAD303"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-Tona"
/clone_lib="potato abiotic stress cDNA library"
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr, 6hr,
12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1d, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:
Pred. No.: 1,23e-54 Length: 929
Score: 881.00 Matches: 194
Percent Similarity: 76.41% Conservatives: 23
Best Local Similarity: 68.31% Mismatches: 53
Query Match: 60.72% Indels: 14
DB: 14 Gaps: 7

US-10-014-927-19 (1-279) x CK272077 (1-929)

QY 4 ArgTrpAsnArgThrileTyrValGlyAsnLeuProGlyAspIleArgLysCysGluVal 23
DB 47 CGACTAAGTCGGACTATCTACGTCGGAATCTTCTGTCATATTCGGAGAGAGAAGTA 106
QY 24 GluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProProArg 43
DB 107 GAGATTGTTTACAAAGTATGTCCTCCATTTGTGAAAATGATTTGAAAGTCTCTCTAGA 166
QY 44 ProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyr 63
DB 167 CCACCTGCTTATGCGTTTGTAGAGTTTGAAGATCTCTGTCGATGCTGATGCTGCTGTC 226
QY 64 GlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGly 83
DB 227 GGGCGTATGCTATGACTTTGACGGCATCGCTTGGAGTTGAACTTGCACATGGTGGG 286
QY 84 ArgArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSerArgAlaPro 103
DB 287 CGAGGA---TCATCATCATATGATCGCACAGTAGTTACAGTAGTGGAGTCGTGGTGA 343
QY 104 ---SerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrp 122
DB 344 CTTTCTAGCGCTCTGACTATCGGTACTGCTCTCTGACTTACCATCTTCTGCTTCATGG 403

QY 200 TyrArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerSerSerArgSer 219
Db 659 ---AGAGTTACTCAAGAACGAGGAGTCTCTCGAACGCTATAGCAGCAGCGGAAGT 715
QY 220 ValSerProAlaArgSerileSerProArgSerArgProLeuSerArgSerSerLeu 239
Db 716 GTATCTCTAGGGTAATAACTCTCGTGGCTCTGTCTGTCTACCTTCAAGGAT--- 772
QY 240 TyrSerSerValSerArgSerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSer 259
Db 773 -----TTTCCCTGCT 784
QY 260 ArgSerIysSerArgSerArgSerArgSerArgSerProValSerPro 275
Db 785 CCCTCTCATTAAGATCTCTATCAAGATCCAGATCCACTTCCACCT 832
RESULT 5
CK244849 935 bp mRNA linear EST 12-DEC-2003
LOCUS EST728486 potato callus cDNA library, normalized and full-length
DEFINITION Solanum tuberosum cDNA clone POC4A66 5' end, mRNA sequence.
ACCESSION CK244849
VERSION CK244849.1 GI:39790831
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 935)
AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE Generation of ESTs from potato callus tissue
JOURNAL Unpublished (2003)
COMMENT Other ESTs: EST728485 EST728487
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers
1..935
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCAA66"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cDNA library, normalized and full-length"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Solanum tuberosum var. Kennebec callus tissue grown on solid media."

ORIGIN
Alignment Scores:
Pred. No.: 1.97e-52 Length: 935
Score: 851.00 Matches: 183
Percent Similarity: 75.19% Conservative: 20
Best Local Similarity: 67.78% Mismatches: 51
Query Match: 58.65% Indels: 16
DB: 14 Gaps: 6
US-10-014-927-19 (1-279) x CK244849 (1-935)

QY 4 ArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGluVal 23
Db 13 CGTCTAAGTCGGAATCTACGTCGGAATCTCTCGTGGTGTATTCGGGAGAGAGTA 72
QY 24 GluAspLeuPheTyrIleTyrGlyProIleValAspIleAspLeuLysIleProArg 43
Db 73 GAAGATTGTTTACAAAGATGATGTCCTCCATGTGGAAATGATTTGAAAGATTCACCTAGA 132

QY 44 ProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIleTyr 63
Db 133 CCACCTGGTTATCGTTCGTAGAGTTTGAAGATCCCTCGTAGTCTGATGATCGCATCCGT 192
QY 64 GlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGlyGly 83
Db 193 GGGCGTGATGGCTATGACTTTGATGGCATCGCTTGGAGTTGAATTCGCACATGTTGGG 252
QY 84 ArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSerArgAlaPro 103
Db 253 CGAGGA---TCATCATCATATGATGCCACAGTAGTTACAGTAGTGGAGTGGTGTGA 309
QY 104 ---SerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAlaSerTrp 122
Db 310 TTTTCTAGGGCTCTGACTATCGCTACTGGTCTCTGGACTACCATCTCTCTGCTTCATGG 369
QY 123 GluAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluValPhePro 142
Db 370 CAAGACTTGAAGGATCACATGCGACAGCTGGAGATGTCTGCTTCTCTCAAGTTTCCGA 429
QY 143 AspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMetLysTyrAla 162
Db 430 GATCGTGATGGTATGAGAGGATGTGGACTATACCACTATGATGATGATGATATGCG 489
QY 163 IleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyrIleArgVal 182
Db 490 ATAAAGAACTTCTGACTCTCTGTTTTCGCAATCAATTCTCTCGAGCATATATTAGGGTG 549
QY 183 ArgGluTyrGlnSerArg---SerValSerArgSerProAspAsp-----SerLysSer 199
Db 550 GACAAGTATGATAAGAGGATGATCTATTCCAGAGTCCCAAGTCCATATTATTTCAGAAGC 609
QY 200 TyrArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSerArgSer 219
Db 610 ---AGAAGTTACTCAAGACAGAGAGTCTTCGACGAGCTATAGCAGCCAGACGGAGT 666
QY 220 ValSerProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArgSerLeu 239
Db 667 GTATCTCTAGGGGTAATAACTCTCGTGGCTGTCTGTCTGTCTCACCTTCAGGGATTTT 726
QY 240 TyrSerSerValSerArgSerGlySerLeuLeuArgAla----- 252
Db 727 TCCTCTGCTCGCTCTCATTCAGATCTGGCTTCGGGGGATCTGGGAATTTGCTATATAG 786
QY 253 -----GlyAspTrpIleSerGlnSer 259
Db 787 ACAGTGTGATGGTTAATGGTCTCAAAATTC 816
RESULT 6
AY108591
LOCUS AY108591 1255 bp mRNA linear HTC 17-OCT-2002
DEFINITION Zea mays P00125156 mRNA sequence.
ACCESSION AY108591
VERSION AY108591.1 GI:21211708
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 1255)
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 2 (bases 1 to 1255)
AUTHORS Coe,E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones,

these are publicly available from ZmDB and may be found by BLAST searching at MGI, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES

Location/Qualifiers
 1..1255
 /organism="Zea mays"
 /mol_type="mRNA"
 /db_xref="taxon:637293"
 /db_xref="taxon:4577"
 /clone_lib="Maize Mapping Project/DuPont Cornsensus Library"
 /note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Alignment Scores:
 Pred. No.: 8,19e-52 Length: 1255
 Score: 845.00 Matches: 186
 Percent Similarity: 73.05% Conservative: 20
 Best Local Similarity: 65.96% Mismatches: 44
 Query Match: 58.24% Indels: 32
 DB: 11 Gaps: 8

US-10-014-927-19 (1-279) x AY108591 (1-1255)

QY 1 MetSerArgTTPAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20
 |||||
 DB 126 ATGACGCGCTGGAGCGCAGCATCTACGTCGGGAACCTCCCGCGACATCCGGGAG 185
 |||||
 QY 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40
 |||||
 DB 186 AGGAGGGTGGAGGATCTCTTCTACAGTAAGTAAATTTGTGACATGACCTGAGGTC 245
 |||||
 QY 41 ProProArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAsp 60
 |||||
 DB 246 CCCCCAAGACCCCTGGTTATGCTTTTCTGAGTTTGAAGATCCTCGTATGCTGAGGAG 305
 |||||
 QY 61 AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80
 |||||
 DB 306 GCAATTTCGGAGCGGATGATACAACTTGTATGACACCGCTCTAAGAGTGGAGGCTGCT 365
 |||||
 QY 81 HisGlyGlyArgArgPheSerProSerValAspArg-----TyrSerSerSerTyr 97
 |||||
 DB 366 CATGCTGGTAGAGGTAATGCTTCTCTCGCATGATCGTTCAAGTGGCTTTGTGGCGGGT 425
 |||||
 QY 98 SerAlaSerArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuPro 117
 |||||
 DB 426 GGAGACGTGTGGTGTGTGGAGACACTCAGAGTATCGTGTCTTGTCACCTGAGCTGCT 485
 |||||
 QY 118 ProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPhe 137
 |||||
 DB 486 TCTTCTGTCATCATGCGAGGATTAAGGATCATATGCGAAGGCTGGTGATGTTTGTTC 545
 |||||
 QY 138 SerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAsp 157
 |||||
 DB 546 TCTGAAGTGTATCGCGAAGCGGTGGCCAGCTAGGATTTGTGGACTACACAAATATGAT 605
 |||||
 QY 158 AspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSer 177
 |||||
 DB 606 GATATGAATATGCTATAAAGAGCTGGATGATATCTGAATTCAGGAAGCCCTTTGGCGGA 665
 |||||
 QY 178 AlaTyrIleArgValArgGluTyrGluSerValSerArgSerValSerArgSerProAspAsp 197
 |||||
 DB 666 GCTATATTAAGGTGAGGAATAT-----AACGGC 695
 |||||
 QY 198 LysSerTyrArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSer 217
 |||||

DB 696 AAACGTGGTGGTCTCTACTCTACGTAGCGGAGCCCAAGTCGTAGTTAC---AGCAAAAGC 752
 |||||
 QY 218 ArgSerValSerProAlaArgSerIleSerProArgSerArg-----ProLeuSer 234
 |||||
 DB 753 AGGAGTCCGAGTAAA-----TCACCCAGGACTCGCGCTTCATCATCTAGATCC 800
 |||||
 QY 235 ArgSerArgSerLeuTyrSerSerValSerArgSerGlySerLeuLeuArgAlaGlyAsp 254
 |||||
 DB 801 CGGTCAAGATCTGTT---TCTTCTCGTTCAAGGTCCCA----- 836
 |||||
 QY 255 TrpIleSerGlnSerArgSerLysSerArgSerArgSerArgSerAsnSerPro---Val 273
 |||||
 DB 837 -----TCAAAGACGCTTCTCCATCAAGATCACCAGCAAGATCGAAATCTCTTAATGTT 890
 |||||
 QY 274 SerPro 275
 |||||
 DB 891 TCTCCA 896

RESULT 7

CG848083

LOCUS

DEFINITION

CG848083

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

CG848083

ZmBBB0319124.1 GI:38374944

Zea mays subsp. mays (maize)

Zea mays subsp. mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 864)

Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J.

Sequencing of the maize genome

Unpublished (2003)

Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: T7

BACKWARD: M13r

Plate: 0319 row: 1 column: 24

Seq primer: M13r

Class: BAC ends

Location/Qualifiers

1..864

/organism="Zea mays subsp. mays"

/mol_type="genomic DNA"

/cultivar="B73"

/sub_species="mays"

/db_xref="taxon:4578"

/clone="ZmBBB0319124"

/lab_host="DH10B"

/clone_lib="ZmBBB"

/note="Vector: pBelOAC11; Site 1: HindIII; Site 2:

HindIII; Zea mays L. ssp. mays"

Alignment Scores:

Pred. No.: 1.04e-51 Length: 864

Score: 840.50 Matches: 175

Percent Similarity: 79.20% Conservative: 23

Best Local Similarity: 70.00% Mismatches: 37

Query Match: 57.93% Indels: 15

DB: 29 Gaps: 5

US-10-014-927-19 (1-279) x CG848083 (1-864)

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

QY	21	CysGluValGluAapLeuPheTyrLysTyrGlyProIleValAspIleAspIleuLysIle	40
DB	163	AGGGAAGTGGGAAGATTGTGTTTAAAGTATGGTCTATAGCTCAAAATTGACTGGAAGATT	222
QY	41	ProProArgPProProGlyTyAlaPheValGluPheGluAapProArgAspAlaAspAsp	60
DB	223	CCACCAAGCCCTCAGGTTATGCATTTGTGTAGTTTGAAGAAGCTCGAGATGCTGAAGAT	282
QY	61	AlaIleTyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla	80
DB	283	GCCATTCTGTGTCGTATGAGTATGACTTTGGAGGGCATCGTTACGGGTGAACCTGTCT	342
QY	81	HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSer---	99
DB	343	CATGGTGTGTGGGGCGCTCA---TCCATAGATCGTCACAGCAGTTATAGTAGTGGCGGT	399
QY	100	SerArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSer	119
DB	400	GGACGTGGACCATCCAGCGCGTCTTGATATCCGGTGTAGTTACTGGATGGCATCTTCT	459
QY	120	AlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGlu	139
DB	460	GCTTCATGCCAGGACCTCAAGATCACATGCGTCGAGCAGGAGATGTTGTGTTTCCCAA	519
QY	140	ValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMet	159
DB	520	GTTTTTCGTGATGGTAGTGGGACACACAGGAATTGTGGACTACACCAACTATGATGATG	579
QY	160	LysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyr	179
DB	580	AGATATGCTATTAGAACAATGATGACTCTGAGTTTCGTAATGCATTTCTCGGCAATAT	639
QY	180	IleArgValArgGluTyrGluSer-----	188
DB	640	GTTCCGGTTAAGGAATATGATTCTAGCGGGGATTCTCTCGAAGCCCTAGTCGTGGCCGA	699
QY	199	SerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSerArgGly	208
DB	700	TCTCTCTCAAGAAGCAGAACGCCAAGCCGGAAGT---AGAAGCCGAAGACGAAGCCGTGT	756
QY	209	ProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSerIleSerPro	228
DB	757	CGAGCTATAGC---CGAAGCAAGAGCGCAGCAGCAAGTCTCCAAAGGCCAAGCCCTCGCG	813
QY	229	ArgSerArgProLeuSerArgSerArgSer	238
DB	814	AAGTCACCTGAAAGCTGTAGATCAAGGTCT	843

RESULT 8	
CK267541	
LOCUS	877 bp mRNA linear EST 12-DEC-2003
DEFINITION	EST173619 potato abiotic stress cDNA library Solanum tuberosum cDNA clone POACH28 5' end, mRNA sequence.
ACCESSION	CK267541
VERSION	CK267541.1
KEYWORDS	GI:39824519
SOURCE	EST.
ORGANISM	Solanum tuberosum (potato)
	Solanum tuberosum
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE	1 (bases 1 to 877)
AUTHORS	Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
TITLE	Generation of ESTs from abiotic stressed potato tissue
JOURNAL	Unpublished (2003)
COMMENT	Contact: Robin Buell The Institute for Genomic Research

FEATURES
SOURCE

```

1. 877
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POACB28"
/tissue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-TonA"
/clone_lib="potato abiotic stress cDNA library"
/note=Vector: PCMVSPORT6.1; Site_1: EcoRI; Site_2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots:2hr, 6hr, 12hr, and 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots:3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:
Pred. No.: 1.37e-51 Length: 877
Score: 839.00 Matches: 185
Percent Similarity: 72.63% Conservatives: 22
Best Local Similarity: 64.91% Mismatches: 48
Query Match: 57.82% Indels: 30
DB: 14 Gaps: 7

```

ORIGIN

```

Alignment Scores:
Pred. No.:      Length:      877
Score:          Matches:     185
              839.00
Percent Similarity: 72.63%
Best Local Similarity: 64.91%
Query Match:       57.82%
                  Indels:    30
                  Gaps:      7
                  DB:         14

US-10-014-927-19 (1-279) x CK267541 (1-877)

QY   3 SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgyysCysGlu 22
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db   23 AGTCGTGCAGTAGACACTTTATGTGGCAATCTTCTGTGTGATGTTCTGTAGCGTGAA 82

QY   23 ValGluAspLeuPheTyrIlystYrGlyProLleValAspIleAspHeuylleProPro 42
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db   83 GTGGGAAGATCTGTTCCCAAGATGCCCCGATAGTCTATTGAGCTGAAAATTCACC 142

QY   43 ArgProProGlyTyralAlaPheValGluPheGluAspProArgAspAlaAspAlaile 62
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db   143 AGACCCCTCGTTATGCTTTTGTTCAGTTTTGAAGAGGCTCGCATGCTCAAGATGCTAIT 202

QY   63 TyrGlyArgAspGlyTyrraspPheaspGlyCysArgLeuArgValGlullealaHisGly 82
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db   203 CGTGTGTCGTATGGGTATGATTTTATGGGCCACCTCTCGCGGTGAGTTGTCACATGGC 262

QY   83 GlyArgArgPheSerProSerValAspArgtyrSerSerSeryrSerAla ----- 99
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db   263 GGCGTGGTAATCATCATCAATAATGATCGTTATGTTGGCGCGGTGTCGCGCGCGC 322

QY   100 -----SerArgAlaProSerArgArgASerAspTyArgValIeuValThrGly 115
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db   323 CQTGTGTCAGCGTGGTGGAGAGTGTCCACACGCTTCTGATTTTCGAGTGTAGTTACAGGA 382

```

116 LeuProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspVal 135
 383 TTACCCCAATCAGCATCGTGGCAGGATCTCAAGGATCAGATGCGTGGGATGTT 442

136 CysPheSerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsn 155
 443 TGTTCCTCCCAAGTTTTCGTCGATGGAGTGGCAGCCACTGGTATAAATGATATACAAAC 502

156 TyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheAsnAlaPhe 175
 503 TATGATGACATGAATATCTCTCAAAAGCTTGATGACTCTGAGTTTCGGAATGCCTTT 562

176 SerSerAlaTyrIleArgValArgGluTyrGlu---SerArgSerValSerArgSerPro 194
 563 TCTCGTCGACAAATTCGAGTTAAGGAATATGACCGCAGTCGAGCCGAGCGCGC 622

195 AspAspSerLysSerTyr-----ArgSerArgSerArgSerArg 207
 623 AGTCGAGCGCTTCTTACTCCAAAGAAAGAGTGTTAGTCGAAGCGGTAGCGTAGCCGA 682

208 GlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSerIleSer 227
 683 AGTAGAAGT-----CGTAGCAAGAGCAGCAAAATCTCTTAAGTTAAGTCTTCA 733

228 ProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerArgSerGly 247
 734 AAGCGCTCAAGA-----TCTGTTCAAGGCTCTG---TCTTCTCGTCTCTGTTCTGGG 784

248 SerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSerLysSerArgSerArgSer 267
 785 -----TCAAAAGAGCGTCTCTGTCTCAAGATCTCCATCA 817

268 ArgSerAsnSerPro 272
 818 AGGTCCAGATCGCCA 832

RESULT 9
 CK250556
 LOCUS
 DEFINITION
 EST734193 potato callus cDNA library, linear EST 12-DEC-2003
 Solanum tuberosum cDNA clone FOCB90 5' end, mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Other ESTs: EST734194
 Contact: Robin Buell
 The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA
 Email: potato-array@tigr.org
 Clones can be requested from TIGR via potato@tigr.org
 Seq primer: ANT TAG GTG ACA CTA TAG.

FEATURES
 source
 1. .910
 /organism="Solanum tuberosum"
 /mol_type="mRNA"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="FOCB90"
 /tissue_type="callus"
 /lab_host="DH10B-TonA"
 /clone_lib="potato callus cDNA library, normalized and full-length"
 /note="Vector: pCMVSPORT6.1; Site_1: EcoRI, Site_2: NotI; supplier: RNA was isolated from Solanum tuberosum var.

Kennebec callus tissue grown on solid media."

ORIGIN

Alignment Scores:

Pred. No.: 1.44e-51 Length: 910
 Score: 839.00 Matches: 185
 Percent Similarity: 72.63% Conservative: 22
 Best Local Similarity: 64.91% Mismatches: 48
 Query Match: 57.82% Indels: 30
 DB: 14 Gaps: 7

US-10-014-927-19 (1-279) x CK250556 (1-910)

Qy 3 SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu 22
 Db 23 AGTCGTCGAGTAGACACTTTATTTGGCAATCTTCTGTGATGTTCTGTGAGCGTGAA 82

Qy 23 ValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProPro 42
 Db 83 GTGGAAGATCTGTTTCAAGTATGCGCCGATAGTCTCATATTGAGCTGAANAATCCACCA 142

Qy 43 ArgProProGlyTyrAlaPheValGluPheValGluAspProArgAspAlaAspAlaIle 62
 Db 143 AGACCCCTGTTTATGCTTTTGTGAGTTTGAAGAGGCTCGCATGCTGAAGATGCTATT 202

Qy 63 TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 82
 Db 203 CGTGTCTGTGATGCTATGATTTGATGGGACCGTCTGCGGTTGAGCTTGACATGGC 262

Qy 83 GlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAla----- 99
 Db 263 GGGCGTGAATCTCATCATCAATGATCGTTATGTTGGCGCGCGGTGGTGGCGGCGC 322

Qy 100 -----SerArgAlaProSerArgArgSerAspTyrArgValLeuValThrGly 115
 Db 323 CGTGTCTAGCGTGGTGAGGAGTGTCCAGACGTTCTGATTTTCGAGTGTAGTTACAGGA 382

Qy 116 LeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspVal 135
 Db 383 TTACCCCAATCAGCATCGTGGCAGGATCTCAAGATCAGATGCTGCTGCTGGGATGTT 442

Qy 136 CysPheSerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsn 155
 Db 443 TGTTCCTCCCAAGTTTTCGTCGATGGAGTGGCAGCCACTGGTATAAATGATTATACAAAC 502

Qy 156 TyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPhe 175
 Db 503 TATGATGACATGAATATGCTATCAAAAGCTTGATGACTCTGAGTTTCGGAATGCCTTT 562

Qy 176 SerSerAlaTyrIleArgValArgGluTyrGlu---SerArgSerValSerArgSerPro 194
 Db 563 TCTCGTCGACAAATTCGAGTTAAGGAATATGACCGCAGTCGAAGCGCAGCCGCGC 622

Qy 195 AspAspSerLysSerTyr-----ArgSerArgSerArgSerArg 207
 Db 623 AGTCGAGCGCTTCTTACTCCAAAGAAAGAGTGTAGTTCGAAGCGGTAGCGTAGCCGA 682

Qy 208 GlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSerIleSer 227
 Db 683 AGTAGAAGT-----CGTAGCAAGAGCAGCAAAATCTCTTAAGTTAAGTCTTCA 733

Qy 228 ProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerArgSerGly 247
 Db 734 AAGCGCTCAAGA-----TCTGTTCAAGGCTCTG---TCTTCTCGGCTCTGTTCTGGG 784

Qy 248 SerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSerLysSerArgSerArgSer 267
 Db 785 -----TCAAAAGAGCGTCTCTGTCTCAAGATCTCCATCA 817

Qy 268 ArgSerAsnSerPro 272
 Db 818 AGGTCCAGATCGCCA 832

```
RESULT 10
CK256686      921 bp      mRNA      linear      EST 12-DEC-2003
LOCUS      EST740323 potato callus cDNA library, normalized and full-length
DEFINITION      Solanum tuberosum cDNA clone POCV33 5' end, mRNA sequence.
ACCESSION      CK256686
VERSION      CK256686.1 GI:39813666
KEYWORDS      EST.
SOURCE      Solanum tuberosum (potato)
ORGANISM      Solanum tuberosum
REFERENCE      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE      asterids; lamids; Solanales; Solanaceae; Solanum.
JOURNAL      1 (bases 1 to 921)
COMMENT      Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
              Generation of ESTs from potato callus tissue
              Other ESTs: EST740324
              Contact: Robin Buell
              The Institute for Genomic Research
              9712 Medical Center Dr, Rockville, MD 20850, USA
              Email: potato-array@tigr.org
              Clones can be requested from TIGR via potato@tigr.org
              Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES
    source
    Location/Qualifiers
        1..921
            /organism="Solanum tuberosum"
            /mol_type="mRNA"
            /cultivar="Kennebec"
            /db_xref="taxon:4113"
            /clone="POCV33"
            /tissue_type="callus"
            /lab_host="DH10B-TonA"
            /clone_lib="potato callus cDNA library, normalized and
            full-length"
            /notes="Vector: pCMVSPORT6.1; Site1: EcoRI; Site 2: NotI;
            supplier: RNA was isolated from Solanum tuberosum var.
            Kennebec callus tissue grown on solid media."
ORIGIN
Alignment Scores:
Pred. No.:      1,47e-51      Length:      921
Score:      839.00      Matches:      185
Percent Similarity:      72.63%      Conservative:      22
Best Local Similarity:      64.91%      Mismatches:      48
Query Match:      57.82%      Indels:      30
DB:      14      Gaps:      7
US-10-014-927-19 (1-279) x CK256686 (1-921)
Qy      3      SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspLeuArgLysCysGlu      22
Db      32      AGTCGGTCGAGTAGGACACTTATGTTGGCAATCTTCTGCGTATGTTGAGCGGTGAA      91
Qy      23      ValGluAspLeuPheTyrLysTyrGlyProIleValAspLeuLysIleProPro      42
Db      92      GTGGAAGATCTGTTTCAAGATAGTGGCCGATAGCTCATATGAGCTGAAATTCACCA      151
Qy      43      ArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIle      62
Db      152      AGACCCCTGGTATGCTTTGTTGATGTTGAGAGCGCTCGATGCTGAGATGCTATT      211
Qy      63      TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluLeuAlaHisGly      82
Db      212      CGTGGTCGTATGCTATGATTTTGTATGGCCGCGCTGCGGGTTGAGCTGACATGGC      271
Qy      83      GlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAla-----      99
Db      272      GGGCGTGGTAACATCATCAATCAATGATCGTATTGTTGGCGCGCGCGTGGTGGCGCGGC      331
Qy      100      -----SerArgAlaProSerArgArgSerAspTyrArgValleuValThrGly      115
Db      332      CGTGGTCAGCGTGGTGGAGGAGTGCAGACGTTCTGATTTTCGAGTGTAGTACAGGA      391
```

```
Qy      116      LeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspVal      135
Db      392      TTACCCCATTCAGCATCGTCGAGGATCTCAAGGATCACATGCTGCTGCTGGGAGTGT      451
Qy      136      CysPheSerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsn      155
Db      452      TGTTCCTCCCAAGTTTCCGTGATGGAGTGGCCACCACTGGTATATGATTATACAAAC      511
Qy      156      TyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPhe      175
Db      512      TATGATGACATGAATATGCTATCAAAAGCTTGATGACTCTGAGTTTCGGAATGCCTTT      571
Qy      176      SerSerAlaTyrIleArgValArgGluTyrGlu---SerArgSerValSerArgSerPro      194
Db      572      TCTCGTCGACAAATTCGAGTTAAGAAATATGACCCGAGTCAAGCCGAGCCGAGCGGC      631
Qy      195      AspAspSerLysSerTyr-----ArgSerArgSerArgSerArgSerArg      207
Db      632      AGTCGGAGCGCTTCTTACTCCAAAGGAAAGAGTGTAGTGAAGCCGTAGCCGTAGCGGA      691
Qy      208      GlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSerIleSer      227
Db      692      AGTAGAAGT-----CGTAGCAAGCAGCAGAGCAATCTCTAAAGTTAAGTCTTCA      742
Qy      228      ProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerArgSerGly      247
Db      743      AAGCGCTCAAGA-----TCTCGTTCAAGGTCTCTG---TCTTCTCGGTCTCGTCTGGG      793
Qy      248      SerLeuLeuArgAlaGlyAspTrpIleSerClnSerArgSerLysSerArgSerArgSer      267
Db      794      -----TCAAAAGGAGCGTCTGCTCAAGATCTCCATCA      826
Qy      268      ArgSerAsnSerPro      272
Db      827      AGGTCAGATCGCCA      841
RESULT 11
CK275124
LOCUS      CK275124
DEFINITION      CK275124 896 bp mRNA linear EST 12-DEC-2003
ACCESSION      EST721202 potato abiotic stress cDNA library Solanum tuberosum cDNA
VERSION      CK275124
KEYWORDS      clone P0ADL50 5' end, mRNA sequence.
SOURCE      CK275124.1 GI:39832102
ORGANISM      Solanum tuberosum (potato)
REFERENCE      Solanum tuberosum
AUTHORS      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
TITLE      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
JOURNAL      asterids; lamids; Solanales; Solanaceae; Solanum.
COMMENT      1 (bases 1 to 896)
              Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.
              Generation of ESTs from abiotic stressed potato tissue
              Unpublished (2003)
              Other ESTs: EST721203
              Contact: Robin Buell
              The Institute for Genomic Research
              9712 Medical Center Dr, Rockville, MD 20850, USA
              Email: potato-array@tigr.org
              Clones can be requested from TIGR via potato@tigr.org
              Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES
    source
    Location/Qualifiers
        1..896
            /organism="Solanum tuberosum"
            /mol_type="mRNA"
            /cultivar="Kennebec"
            /db_xref="taxon:4113"
            /clone="P0ADL50"
            /tissue_type="abiotic stress treated leaf and root tissue"
            /lab_host="DH10B-TonA"
            /clone_lib="potato abiotic stress cDNA library"
            /note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
            supplier: Solanum tuberosum var. Kennebec plants were
```

grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Alignment Scores:
Pred. No.: 3,02e-51 Length: 896
Score: 834.50 Matches: 184
Percent Similarity: 72.73% Conservative: 24
Best Local Similarity: 64.34% Mismatches: 49
Query Match: 57.51% Indels: 29
DB: 14 Gaps: 7

US-10-014-927-19 (1-279) x CK275124 (1-896)

QY 3 SerArgTrpAnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu 22
DB 22 AGTCGTCGAGTAGACACTTATGTTGGCAATCTTCCTGGTGAATGCCGAGCGTGA 81
QY 23 ValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIlePro 42
DB 82 GTGGAGATCTCTTTCACAGTATGCGCCGATAGCTCATATTGAGTGAATTCACCA 141
QY 43 ArgProPcGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIle 62
DB 142 AGACCCCTGTTATGCTTTGTTGAGTTTGAAGAGGCTCGTGTGCTGAAGATGCTATT 201
QY 63 TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 82
DB 202 CGTGTGCTGATGCTATGTTTTCATGCGCACCCGCTTCGCGGTGAGCTTGCACATG 261
QY 83 GlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAla----- 99
DB 262 GGGCGTGGTAACTCATCATCAATGATGTTATGTTGGCGCGCGGCGGTGGTGGCGGC 321
QY 100 -----SerArgAlaProSerArgArgSerAspTyrArgValLeuValThr 114
DB 322 GGGCGTGGTACCGTGGTGGAGAGTGTCCAGACGCTTCTGATTTTCGAGTGTATTACA 381
QY 115 GlyLeuProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAsp 134
DB 382 GGATTACCCCATTCAGCATCGTGGCAGGATCTCAAGGATCATCATCGTCTGCTGGGAT 441
QY 135 ValCysPheSerGluValPheProAspArgLysGlyMetSerGlyValValAspTyrSer 154
DB 442 GTTTGTTTCTCCCAAGTTTTCGCTGATGGAGTGCACCCACTGGTATGATTGATTATACA 501
QY 155 AsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAla 174
DB 502 AACTATGATGACATGAAATATGCTATCAAAAGCTTGATGACTCTGAGTTTCGATGCC 561
QY 175 PheSerSerAlaTyrIleArgValArgGluTyrGlu---SerArgSerValSerArgSer 193
DB 562 TTTTCTCGTGGACATTCGAGTTAAGGAATATGACCGCAGTCTGAAGCGCAGCGCAGC 621

QY 194 ProAspAspSerLysSerTyr-----ArgSerArgSerArgSer 206
DB 622 CGCAGTCGAGCCGCTTCTTACTCCAAAGGAAGAGTGTAGTCGAAGCCGTAGCCGTAGC 681
QY 207 ArgGlyProSerCysTyrSerSerLysSerArgSerValSerProAlaArgSerIle 226
DB 682 CGTAGCCGAGCAAAAGT---CGTAGCAAGACGACAAATCTCTAAAGTTAAGTCT 738
QY 227 SerProArgSerArgProLeuSerArgSerArgSerLeuTyrSerValSerArgSer 246
DB 739 TCAAGCGCTCAAGA-----TCTCGTTCAGGTTCTGUG---TCTTCTCGTCTCGTCT 789
QY 247 GlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSerLysSerArgSerArg 266
DB 790 GGG-----TCAAAAGGACGCTCTGTCTCAAGATCTCCA 822
QY 267 SerArgSerAsnSerPro 272
DB 823 TCAAGGTCAGATCGCCA 840

RESULT 12

CG848162

LOCUS

DEFINITION

CG848162

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

CG848162

ZM5BB0319K18.1

GI:38375023

GSS.

Zea mays subsp. mays (maize)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 874)

Yu, Y., Kim H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J. and Wing, R.

Sequencing of the maize genome

Unpublished (2003)

Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: T7

BACKWARD: M13r

Plate: 0319 row: K column: 18

Seq primer: M13r

Class: BAC ends.

Location/Qualifiers

1. .874

/organism="Zea mays subsp. mays"

/mol_type="genomic DNA"

/cultivar="B73"

/sub_species="mays"

/db_xref="taxon:4578"

/clone="ZM5BB0319K18"

/lab_host="DH10B"

/clone_lib="ZM5BBB"

/notes="Vector: pBelBAC11; Site 1: HindIII; Site 2: HindIII; Zea mays L. asp. mays"

ORIGIN

Alignment Scores:

Pred. No.: 6,79e-51 Length: 874

Score: 829.50 Matches: 180

Percent Similarity: 75.09% Conservative: 25

Best Local Similarity: 65.93% Mismatches: 43

Query Match: 57.17% Indels: 25

DB: 29 Gaps: 8

US-10-014-927-19 (1-279) x CG848162 (1-874)

Qy 1 MetSerSerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLys 20

Db 115 ATGAGTAGCGGTCCAGCAGGACTCTTTACGTGTGGAAATCTTCCCGGTGATGTCGCGAG 174

Qy 21 CysGluValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIle 40

Db 175 AGGGAAGTGGAGATTTGTTTATAGTATGTTGCTTATAGTCAAAATTCACATTGAAGATT 234

Qy 41 ProProArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAsp 60

Db 235 CCACCAAGGCTCCAGGTATGATTTGATTTGATTTGAAGAAGTCGAGATGCTGAAGAT 294

Qy 61 AlaIleTyrClyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAla 80

Db 295 GCCATTGCGTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 354

Qy 81 HisGlyGlyArgArgPheSerProSerValAspArgTyrSerSerSerSerTyrSerAla 99

Db 355 CATGGTGGTGGTGGGCGCTCA---TCCATAGATCGTCACAGCATTTATAGTAGTGGCGT 411

Qy 100 SerArgAlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProSer 119

Db 412 GACGCTGGACCATCCAGGCGTTCTGAATATCGCGTCTAGTTACTGATGCTCATCTTCT 471

Qy 120 AlaSerTrpClnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGlu 139

Db 472 GCTTCATGGCAGACCTCAAGGATCATCGCTCGAGCAGGAGATGTTGTTTTCGCCAA 531

Qy 140 ValPheProAspArgLysGlyMetSerGlyValValAspTyrSerAsnTyrAspAspMet 159

Db 532 GTTTTCGTGATGGTAGTGGACTACAGGATTTGGACTACCACTATGATGATGATG 591

Qy 160 LysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsnAlaPheSerSerAlaTyr 179

Db 592 AGTATGCTATTAGAACTGGATGATCTGAGTTTCGTAATGCAATTTCTCGGCAAT 651

Qy 180 IleArgValArgLysGluTyrGluSer---ArgSerValSerArgSerProAspAspSerLys 198

Db 652 GTTCGGGTTAAGGAATATGATTTAGCGCGGATTCCTCTCGAAGCCCTAGTCTGCGCCA 711

Qy 199 SerTyr---ArgSerArgSerArgSerArgGlyProSerCysSerTyrSerSerLysSer 217

Db 712 TCTCTCTCAAGAGCAGACCGACCGA-----AGTAGAAGC 750

Qy 218 ArgSerValSerProAlaArgSerIleSerProArgSerArgProLeuSerArgSerArg 237

Db 751 CGAAGCAGAACCGTGGTGGAGCTAT-----AGCGAAGCAGAACCGCGCAGCAAG 801

Qy 238 SerLeuTyrSerSerValSerArgSerGlySerLeuArgAlaGlyAspTrpIleSer 257

Db 802 TCTCCAAAGCAGAACGCTTCGCGCAAG-----TCA 831

Qy 258 GlnSerArgSerLysSerArgSer---ArgSerArgSer 269

Db 832 CCTGAAAGCTANGATCAGGTCCTCTCGTTCGCGCTCT 870

RESULT 13

CK298330

LOCUS

DEFINITION

EST761044 Nicotiana benthamiana mixed tissue cDNA library,
normalized, full-length Nicotiana benthamiana cDNA clone NBMD19 5',
end, mRNA sequence.

CK298330

CK298330.1 GI:39885597

VERSION

KEYWORDS

SOURCE

ORGANISM

Nicotiana benthamiana

Nicotiana benthamiana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Nicotiana.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 903)

Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B.,
Staskawicz,B., Jin,H. and Baker,B.
Generation of EST sequences from Nicotiana benthamiana
Unpublished (2003)
Other_ESTs: EST761045
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers

1.903

/organism="Nicotiana benthamiana"

/mol_type="mRNA"

/db_xref="taxon:4100"

/clone="NBMD19"

/tissue_type="abiotic and biotic stress-treated leaves,
callus tissue and root tissue"

/lab_host="DH10B-Tona"

/clone_lib="Nicotiana benthamiana mixed tissue cDNA
library, normalized, full-length"

/notes="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Nicotiana benthamiana
tissues that include callus, roots from liquid culture
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen
challenged leaves (Pseudomonas syringae pv tomato 12 hr;
Xanthomonas campestris pv campestris 12 hr, 18hr;
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas
campestris pv vesicatoria 18hr). RNA was isolated from
these tissues and pooled in approximately equal molar
amounts."

Alignment Scores:

Align. No.: 7,1e-51 Length: 903

Score: 829.50 Matches: 185

Percent Similarity: 74.48% Conservative: 28

Best Local Similarity: 64.69% Mismatches: 44

Query Match: 57.17% Indels: 29

DB: 14 Gaps: 9

US-10-014-927-19 (1-279) x CK298330 (1-903)

Qy 3 SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu 22

Db 79 AGTCGATCGATAGGACACATTTATGTTGGCAATCTTCCGGGCGATATTCGTGACGCGAA 138

Qy 23 ValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIlePro 42

Db 139 GTGGAAGATCTGTTTCAAGTATGGCCCAATAGCCGATATTTAGAGCTGAATAATTCACCA 198

Qy 43 ArgProProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIle 62

Db 199 AGGCCCTCGTTGATGCTTTGTTGAGTTGAGAAAGTTCGTGATGCTGAAGATGCTATT 258

Qy 63 TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 82

Db 259 CGTGGTGTGATGCTATGATTTTGTATGGGCACCGCTCTGAGGGTTGAGCTTGCACATGTT 318

Qy 83 GlyArgArgPheSerProSerValAspArgTyrSerSerSerTyrSerAlaSerArg--- 101

Db 319 GGGGCTGTCTAGTCA---TCAAACGATCGTTATGCTGTGTGTGAGGTCGTGGTCAACGT 375

Qy 102 ---AlaProSerArgArgSerAspTyrArgValLeuValThrGlyLeuProProSerAla 120

Db 376 GGTGGAGTGTCCAGACGTTCTGATTATCGAGTGAATAATTCAGGATTTACCCATTCACCG 435

Qy 121 SerTrpGlnAspLeuLysAspHisMetArgLysAlaGlyAspValCysPheSerGluVal 140

Db 436 TCATGGCAGGATCTCAGGATCACAATCGTCGAGCTGGGATGCTGTTTCTCCCAAGTT 495

Db 787 -----TCAAAGGGCGTTCTATCTCAAGATCTCCATCAAGGTCCAGATCGCCA 834
Qy 273 ValSerProValIleSer 278
Db 835 ---GCACCATCTGTAGT 849

RESULT 15
CK265450 884 bp mRNA linear EST 12-DEC-2003
LOCUS EST711528 potato abiotic stress cDNA library Solanum tuberosum CDNA
DEFINITION clone POAB79 5' end, mRNA sequence.
ACCESSION CK265450
VERSION CK265450.1 GI:39822428
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
REFERENCE Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
AUTHORS Generation of ESTs from abiotic stressed potato tissue
TITLE Unpublished (2003)
JOURNAL Other ESTs: EST711529
COMMENT Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES
1. 884
Location/Qualifiers
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POAB79"
/issue_type="abiotic stress treated leaf and root tissue"
/lab_host="DH10B-TonA"
/clone_lib="potato abiotic stress cDNA library"
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: Solanum tuberosum var. Kennebec plants were
grown from cuttings on a 16hr light/8 hr dark cycle at 25
C for 3-4 weeks. Abiotic stress conditions were applied to
four separate sets of plants. Set 1 involved saturation of
the soil with 150 mM NaCl and tissues were harvested at
following application of the salt stress (leaves: 2hr,
6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, 1d, 2d).
Set 2 were grown under the standard conditions and then
were water stressed by withdrawal of further watering
applications. Drought stressed plants were harvested after
cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
and 5d). Set 3 were grown under the standard conditions
and then were cold stressed by placement at 4 C. Cold
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d, 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d. Set 4 were grown under the standard conditions and
then were heat stressed by placement at 35 C. Heat
stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
2d and 4d and heat-stressed roots were harvested at 6 hr,
12 hr, 1 d, and 4d. RNA was isolated from all tissues and
equal RNA from each tissue and stress was pooled to
construct the cDNA library. RNA sample."

ORIGIN
Alignment Scores:
Pred. No.: 7.5e-51 Length: 884
Score: 829.00 Matches: 182
Percent Similarity: 71.78% Conservative: 24
Best Local Similarity: 63.41% Mismatches: 47
Query Match: 57.13% Indels: 34
DB: 14 Gaps: 7

US-10-014-927-19 (1-279) x CK265450 (1-884)
Qy 3 SerArgTrpAsnArgThrIleTyrValGlyAsnLeuProGlyAspIleArgLysCysGlu 22
Db 23 AGTCGTCGAGTAGGACACCTTTATGTTGGCAACTTCTCTGATGATGTTCTGAGCGTGAA 82
Qy 23 ValGluAspLeuPheTyrLysTyrGlyProIleValAspIleAspLeuLysIleProPro 42
Db 83 GTGGAAGATCTGTTTCAAGAATGATGCGCCGATAGTCATATTGAACCTGAAATATCCACCA 142
Qy 43 ArgProGlyTyrAlaPheValGluPheGluAspProArgAspAlaAspAlaIle 62
Db 143 AGACCCCTCGTTATGCTTTTGTGAGTTTCAAGAGCTCGCATGCTGAAGATGCTATT 202
Qy 63 TyrGlyArgAspGlyTyrAspPheAspGlyCysArgLeuArgValGluIleAlaHisGly 82
Db 203 CGTGGTCTGATGCTATGATTGTTGAGGACCGCTCTGCGGTTGAGCTTGCACATGGC 262
Qy 83 GlyArgArgPheSerProSerValAspArgTyrSerSerTyrSerAla----- 99
Db 263 GGGCGGTAACTCATCATCAATGATCGTTATGCGCGCGGTGGTGGTGGTGGC 322
Qy 100 -----SerArgAlaProSerArgArgSerArgTyrArgValLeuVal 113
Db 323 GCGCGCGTGTGTCAGCGTGGTGGAGAGTGTCCAGACGTTCTGATTTTCAGTGTAGTT 382
Qy 114 ThrGlyLeuProProSerAlaSerTrpGlnAspLeuLysAspHisMetArgLysAlaGly 133
Db 383 ACAGGATTACCCCATTCAGCATCTGTCAGAGATCTCAAGGATCACATGCGTCTGCTGGG 442
Qy 134 AspValCysPheSerGluValPheProAspArgLysGlyMetSerGlyValValAspTyr 153
Db 443 GATGTTGTTTCTCCCAAGTTTCCGTGATGGGAGTGGCACCACTGGTATATATGATAT 502
Qy 154 SerAsnTyrAspAspMetLysTyrAlaIleArgLysLeuAspAlaThrGluPheArgAsn 173
Db 503 ACAAACTATGATGACATGAATATGCTATCAAAAGCTTGATGACTGCTGAGTTTCGGAAT 562
Qy 174 AlaPheSerSerAlaTyrIleArgValArgGluTyrGlu---SerArgSerValSerArg 192
Db 563 GCCTTTTCTCGTCGACAAATTCGAGTTAGGAATATGACCGCATGCGAGCCGAGCCGC 622
Qy 193 SerProAspAspSerLysSerTyr-----ArgSerArgSerArg 205
Db 623 AGCGCGAGTCGAGCGCGTCTTCTTATCCAAAGAAAGAGTGTAGTCGAAGCCGTAGCCGA 682
Qy 206 SerArgGlyProSerCysSerTyrSerSerLysSerArgSerValSerProAlaArgSer 225
Db 683 AGTAGAGT-----CGTAGCAAGCAAGCAAAAGCAATCTCTTAAGTTAAG 727
Qy 226 IleSerProArgSerArgProLeuSerArgSerArgSerLeuTyrSerSerValSerArg 245
Db 728 TCTTCAAGGCGCTCAAGA-----TCTCGTTCAAGGTCTGTG--TCTTCTCGTCTCGT 778
Qy 246 SerGlySerLeuLeuArgAlaGlyAspTrpIleSerGlnSerArgSerLysSerArgSer 265
Db 779 TCTGGG-----TCAAAGGACGCTCTCTGTCTCAAGATCT 811
Qy 266 ArgSerArgSerAsnSerPro 272
Db 812 CCATCAAGGTCAGATCGCCA 832

Search completed: March 26, 2004, 20:24:40
Job time : 2837 secs

(02Cb)

1
RESULT 5
ATH131214
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

ATH131214 5164 bp DNA linear PLN 07-MAY-1999
Arabidopsis thaliana srp30 gene, exons 1-12.
AJ131214
AJ131214.1 GI:4775269
SF2/ASF-like splicing modulator; srp30 gene.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 Lopato, S., Kalyana, M., Dörner, S., Kobayashi, R., Krainer, A.R. and Barta, A.
AUTHORS
TITLE atSRP30, one of two SF2/ASF-like proteins from Arabidopsis
JOURNAL thaliana, regulates splicing of specific plant genes
MEDLINE Genes Dev. 13 (8), 987-1001 (1999)
PUBMED 99234087

REFERENCE 2 (bases 1 to 5164)
PUBMED 10215626

AUTHORS Barta, A.
TITLE Direct Submission
JOURNAL Submitted (02-DEC-1998) Barta A., Institute for Biochemistry,
University of Vienna, Dr. Bohrgasse 9/3,, A-1030 Vienna, AUSTRIA

FEATURES
source

1. 5164
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
join(1806..2014,2138..2203,2290..2364,2444..2546,
2632..2678,2786..2843,2968..3017,3161..3217,3291..3403,
3501..3578,4515..4572,4844..5125)
1806..2014
/number=1
1928..4858
/gene="srp30"

mRNA

exon

gene


```
Db 2347 GATGGGTGTGCACTTCGGGTTAGTAAACGATGATGAAAGCTAGCTTAATTTCTGTAA 2406
Qy 77 -----ValGluIleAlaHisGlyLys 84
Db 2407 TTCTTGTAAAGGTGTATCTTCTGTGATGTTTTTAGGTGAGATTGCCACATGGGGTGG 2465
Qy 84 gArgPheSerProSerValAspArgTyrSerSerTyrSerAlaSerArgAlaProSe 104
Db 2467 TAGATTTTACCAATCAGTTGATAGTACAGCAGCAGCTACAGCGCGGTGCACCTTC 2526
Qy 104 rArgArgSerAspTyrArg----- 110
Db 2527 AAGACGCTCTGACTACCGCGTTTGTAGAGTCTTCTCGATTGTGTTATTGGTGTGTGT 2586
Qy 111 -----ValLeuValThrGlyL 116
Db 2587 AAAATTTTATATTGAAAACCTCAATTTTACTACCTAAACATGATGCTTGTGACCGGAT 2646
Qy 116 euProProSerAlaSerTyrGlnAspLeuLys----- 126
Db 2647 TACCGCCTTCTGCTTCGTGGCAGGACCTTAA-GGTAAGGGACACATATATAGTCTTTTCT 2705
Qy 126 ----- 126
Db 2706 CTGATGTGGTCTCTATATATCATGTTTGGATTATCTCTTTCTCGAATGATGTTATT 2765
Qy 127 -----AspHisMetArgLysAlaGlyAspValCysPheSerGluV 140
Db 2766 TGCTATTACGGGTGATTAGGATCACATGCGCAAGCTGGAGATGTCTGCTTCTCTGAAG 2825
Qy 140 alPheProAspArg-Lys----- 145
Db 2826 TTTTCCCTGACCGTAAGGTGAGTTGACATTCGATAGTTTCGATAAGCTTTTGTATTGAT 2885
Qy 145 ----- 145
Db 2886 GGTGTAGTAAATAGTCTTTGTGAAGGAGATAGTGTAAAGCATCTGAACGTGCTAAACT 2945
Qy 146 -----GlyMetSerGlyValValAspTyrSerAsnTyrAspAsp 158
Db 2946 CACATTCAGTATTTCTTTCTAGGCATGCTGGGGTTGTGGATTATAGCAACTATGATGAT 3005
Qy 159 MetLysTyrAla----- 162
Db 3006 ATGAAGTAGCG-AGTAAAGTTTATATCTTTGCAACGCAAAAGTTCCTGGACTTATGCTT 3064
Qy 162 ----- 162
Db 3065 AGACTGCTTTTGTTCATAGTATACGAGCTGAATTTATCTCTGGAGGCGAGTGTGG 3124
Qy 163 -----IleArgLysLeuAspAlaThrGlu 170
Db 3125 ATCTTTGATGTTCCCTTAAATTTTGTGATGTGACAGATAGGAACTTGATGCCACTGAA 3184
Qy 171 PheArgAsnAlaPheSerSerAlaTyrIleArg----- 181
Db 3185 TTTTCGAAATGCTTTCTCTAGTGTATATACGGGTATGTTGTATGCTTTCTTTGATTTT 3244
Qy 182 -----ValArgGluTyrGln 186
Db 3245 GTTAAGCATAAGTGGATATGAGTCACTCTGAAATTTACTGTTTCAGGTGAGGGAATATGA 3304
Qy 186 uSerArgSerValSerArgSerProAspAspSerLysSerTyrArgSerArgSerArgSe 206
Db 3305 GTCGAGGAGTGTGAGTGCAGGCCAGATGATTTCTAAAGCTATAGAAGCAGGAGTCGGAG 3364
Qy 206 rArgGlyProSerCysSerTyrSerSerLysSerArg----- 218
Db 3365 CCGTGGTCCAGCTGTAGTATAGTACCAAGACGAGGAGTCTTTTTTTTTTTTTTTTTTTT 3424
Qy 218 ----- 218
Db 3425 TCATAAACCTAAGACATATAAGGGATTTTATTGTTAACTTATTATGAAATTAACCTGACT 3484
```

```
Qy 219 -----SerValSerProAlaArgSerIleSerProArgSerArgProLeuS 234
Db 3485 TCATAATGCAATGCAATGCTGTGTACCTGTAGATCCATTTCCCGCGGTTTCACGCCCCCTTA 3544
Qy 234 erArgSerArgSerLeuTyrSerSerValSerArg 245
Db 3545 GTCGTTTCTGCTCGCTATATACAGCTCTGTCTCAAGG 3579
```
